

FORM PTO-1399 (Modified)  
(REV 10/95)

U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

ATTORNEY'S DOCKET NUMBER

## TRANSMITTAL LETTER TO THE UNITED STATES

VOSS1100

DESIGNATED/ELECTED OFFICE (DO/EO/US)

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR

CONCERNING A FILING UNDER 35 U.S.C. 371

09/445201

INTERNATIONAL APPLICATION NO.

INTERNATIONAL FILING DATE

PRIORITY DATE CLAIMED

PCT/EP98/03318

June 3, 1998

June 3, 1997

## TITLE OF INVENTION

REGULATORY SEQUENCES CAPABLE OF CONFERRING EXPRESSION OF A HETEROLOGOUS DNA SEQUENCE IN ENDOTHELIAL CELLS IN VIVO AND USES THEREOF

APPLICANT(S) FOR DO/EO/US

GEORG BREIER; WERNER RISAU; VOLKER RONICKE

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371 (c) (2))
  - a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
  - b. ☒ has been transmitted by the International Bureau.
  - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ☐ A copy of the International Search Report (PCT/ISA/210).
8. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
  - a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
  - b. ☐ have been transmitted by the International Bureau.
  - c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
  - d. ☐ have not been made and will not be made.
9. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
10. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).
11. ☐ A copy of the International Preliminary Examination Report (PCT/IPEA/409).
12. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).

## Items 13 to 18 below concern document(s) or information included:

13. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
14. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
15. ☒ A **FIRST** preliminary amendment.  
A **SECOND** or **SUBSEQUENT** preliminary amendment.
16. ☐ A substitute specification.
17. ☐ A change of power of attorney and/or address letter.
18. ☒ Certificate of Mailing by Express Mail EL360645103US
19. ☒ Other items or information:

Postcard

097/445201

514 Rec'd PCT/PTO 03 DEC 1999

20. The following fees are submitted:

## BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :

- ☐ Search Report has been prepared by the EPO or JPO ..... \$840.00
- ☐ International preliminary examination fee paid to USPTO (37 CFR 1.482) ..... \$670.00
- ☐ No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2)) ..... \$760.00
- ☒ Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO ..... \$970.00
- ☐ International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4) ..... \$96.00

## ENTER APPROPRIATE BASIC FEE AMOUNT =

## CALCULATIONS PTO USE ONLY

Surcharge of \$130.00 for furnishing the oath or declaration later than months from the earliest claimed priority date (37 CFR 1.492 (e)).

☐ 20 ☐ 30

\$970.00

\$0.00

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	
Total claims	52 - 20 =	32	x \$18.00	\$576.00
Independent claims	1 - 3 =	0	x \$78.00	\$0.00
Multiple Dependent Claims (check if applicable)			<input type="checkbox"/>	\$0.00

TOTAL OF ABOVE CALCULATIONS =

\$1,546.00

Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28) (check if applicable).

☒

\$773.00

SUBTOTAL =

\$773.00

Processing fee of \$130.00 for furnishing the English translation later than months from the earliest claimed priority date (37 CFR 1.492 (f)).

☐ 20 ☐ 30

\$0.00

TOTAL NATIONAL FEE =

\$773.00

Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) (check if applicable).

☐

\$0.00

TOTAL FEES ENCLOSED =

\$773.00

Amount to be:  
refunded  
charged

\$

\$

- ☐ A check in the amount of \_\_\_\_\_ to cover the above fees is enclosed.
- ☒ Please charge my Deposit Account No. **07-1895** in the amount of **\$773.00** to cover the above fees.  
A duplicate copy of this sheet is enclosed.
- ☒ The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. **07-1895** A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Lisa A. Haile  
Gray Cary Ware & Friederich LLP  
4365 Executive Drive, Suite 1600  
San Diego, California 92121-2189

Phone: (858) 677-1456

Facsimile (858) 677-1465

SIGNATURE

Lisa A. Haile

NAME

38,347

REGISTRATION NUMBER

December 2, 1999

DATE

PATENT

Attorney Docket No.: VOSS1110

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:  
Georg Breier, et al.

) Group Art Unit: Unassigned

) Examiner: Unassigned

Application No.: Unassigned  
Filed: Herewith based on PCT/EP98/03318

) CERTIFICATE OF MAILING BY "EXPRESS MAIL"

) "EXPRESS MAIL" MAILING LABEL NO. EL360645103US

For: REGULATORY SEQUENCES CAPABLE  
OF CONFERRING EXPRESSION OF A  
HETEROLOGOUS DNA SEQUENCE IN  
ENDOTHELIAL CELLS IN VIVO AND USES  
THEREOF

) DATE OF DEPOSIT December 2, 1999. I HEREBY CERTIFY THAT  
) THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES  
) POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE  
) UNDER 37 C.F.R. 1.10 ON THE DATE INDICATED ABOVE AND IS  
) ADDRESSED TO THE ASSISTANT COMMISSIONER FOR PATENTS,  
) WASHINGTON, D.C. 20231

) Michael Bayley

) (TYPED OR PRINTED NAME OF PERSON MAILING PAPER OR FEE)

) *Michael Bayley*  
) \_\_\_\_\_  
) SIGNATURE OF PERSON MAILING PAPER OR FEE)

Assistant Commissioner for Patents  
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Prior to examination of the above-identified new patent application, please  
consider the following amendments and remarks:

**IN THE CLAIMS:**

Please amend the claims as follows:

In claim 4, line 1, please delete “any one of claims 1 to 3” and substitute therefor --claim 1--.

In claim 6, line 1, please delete “or 5”.

In claim 7, line 1, please delete “or 6”.

In claim 9, line 1, please delete “or 8”.

In claim 10, line 1, please delete “to 9”.

In claim 11, line 1, please delete “to 10”.

In claim 12, line 1, please delete “to 11”.

In claim 16, line 3, please delete “to 15”.

In claim 17, line 1, please delete “to 15”.

In claim 19, line 1, please delete “or 18”.

In claim 20, please delete “to 15” and “to 19”.

In claim 22, please delete “or 21”.

In claim 23, please delete “to 22”.

In claim 24, please delete “to 15” and “to 19”.

In claim 25, please delete “to 15”, “to 19” and “to 23”.

In claim 26, please delete “to 15” and “to 19”.

In claim 27, please delete “to 15” and “to 19”.

In claim 28, please delete “or the transgenic non-human animal of claim 27”.

In claim 29, line 4, please delete “to 23” and in line 5, please delete “or 28”.

In claim 30, line 4, please delete “to 23” and in line 5, please delete “or 28”.

In claim 31, line 1, please delete “to 15”; in line 2, please delete “to 19” and “to 23” and in line 4, please delete “or 28”.

In claim 32, line 1, please delete “or 30 or the use of claim 31”.

In claim 33, line 2, please delete “or 30”.

In claim 34, line 2, please delete “to 19”.

In claim 36, line 1, please delete “to 15”; line 2, please delete “to 19” and in line 3, please delete “30 or 32”.

In claim 37, line 1, please delete “to 15”; line 2, please delete “to 19” and in line 3, please delete “30 or 32”.

In claim 38, line 1, please delete "to 15"; line 2, please delete "to 19" and in line 4, please delete "or 28".

In claim 39, line 1, please delete "or 35" and "to 38".

Please add the following new claim:

--41. The transgenic non-human animal of claim 27, where said animal is a mouse.--

#### REMARKS

The present amendments were required to reduce the number of claims for filing the application. No new matter has been added. Applicant respectfully requests entry of this amendment.

If the Examiner would like to discuss any of the issues raised in the Office Action, Applicant's representative can be reached at (858) 677-1456. While it is not believed that any fees are due, if the PTO finds any additional fees due, or make any credits, to Deposit Account No. 07-1895.

Respectfully submitted,



Lisa A. Haile, Ph.D.

Reg. No. 38,347

Telephone: (858) 677-1456

Facsimile: (858) 677-1465

Date: December 2, 1999

GRAY CARY WARE & FREIDENRICH LLP  
4365 Executive Drive, Suite 1600  
San Diego, California 92121-2189

Attorney Docket No.: VOSS1110

Applicant: Breier et al.

Serial No.: 09/445,201

Filed: December 2, 1999

Title: REGULATORY SEQUENCES CAPABLE OF CONFERRING EXPRESSION OF A HETEROLOGOUS DNA SEQUENCE IN ENDOTHELIAL CELLS IN VIVO AND USES THEREOF

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS  
37 C.F.R. §1.9(f) and 1.27(d) - NONPROFIT ORGANIZATION

I hereby declare that I am an official empowered to act on behalf of the nonprofit organization identified below:

NAME OF ORGANIZATION: Maz-Planck-Gesellschaft zur Förderung der Wissenschaften e.v.  
ADDRESS OF ORGANIZATION: Berlin, GERMANY

TYPE OF ORGANIZATION

- ☐ University or other institution of Higher Education  
☒ Tax Exempt under Internal Revenue Service Code (26 U.S.C. §501(c) and 501(c) (3))  
☐ Nonprofit Scientific or Educational under Statute of State of the United States of America (Name of State \_\_\_\_\_) (Citation of Statute \_\_\_\_\_)  
☐ Would qualify as tax exempt under Internal Revenue Service Code (26 U.S.C. §501(a) and 501(c) (3)) if located in the United States of America  
☐ Would qualify as nonprofit Scientific or Educational under Statute of State of the United States of America if located in the United States of America (Name of State \_\_\_\_\_) (Citation of Statute \_\_\_\_\_)

I hereby declare that the nonprofit organization identified above qualifies as a nonprofit organization as defined in 37 C.F.R. §1.9(e) for purposes of paying reduced fees under Section 41(a) and (b) of Title 35, United States Code, with regard to the invention entitled REGULATORY SEQUENCES CAPABLE OF CONFERRING EXPRESSION OF A HETEROLOGOUS DNA SEQUENCE IN ENDOTHELIAL CELLS IN VIVO AND USES THEREOF by inventors Georg Breier; Werner Risau; and Volker Ronke as described in

- ☐ the specification filed herewith  
☒ application Serial No. 09/445,201, filed December 2, 1999  
☐ Patent No. \_\_\_\_\_, issued \_\_\_\_\_

I authorize and request insertion of the serial number of the application when officially known.

I hereby declare that rights under contract or law have been conveyed to and remain with the nonprofit organization with regard to the above-identified invention.

If the rights held by the nonprofit organization are not exclusive, each individual, concern or organization having rights to the invention is listed below\* and no rights to the invention are held by any person, other than the inventor, who could not qualify as a small business concern under 37 C.F.R. §1.9(d) or by any concern which would not qualify as a small business concern under 37 C.F.R. §1.9(d) or a nonprofit organization under 37 C.F.R. §1.9(e).

NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities (37 C.F.R. §1.27).

Full Name \_\_\_\_\_  
Address \_\_\_\_\_  
☐ Individual ☐ Small Business Concern ☐ Nonprofit Organization  
Full Name \_\_\_\_\_  
Address \_\_\_\_\_  
☐ Individual ☐ Small Business Concern ☐ Nonprofit Organization

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate (37 C.F.R. §1.28(b)).

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING: Christa Herzog  
TITLE IN ORGANIZATION: Head of Patent Department  
ADDRESS OF PERSON SIGNING: Hofgartenstr. 8, 80539 München, GERMANY  
SIGNATURE: *Christa Herzog* DATE: 3.4.2000  
Christa Herzog

Rec'd PCT/PTO 02 DEC 1999

REGULATORY SEQUENCES CAPABLE OF CONFERRING EXPRESSION OF A  
HETEROLOGOUS DNA SEQUENCE IN ENDOTHELIAL CELLS IN VIVO AND  
USES THEREOF

The present invention relates to recombinant DNA molecules comprising the regulatory sequence(s) of an intron of the Vascular Endothelial Growth Factor (VEGF) receptor-2 gene (Flk-1) or of a gene homologous to the Flk-1 gene, being capable of conferring expression of a heterologous DNA sequence in endothelial cells in vivo. The present invention also relates to vectors comprising said recombinant DNA molecules as well as to host cells transformed with such recombinant DNA molecules or vectors. The present invention additionally relates to pharmaceutical and diagnostic compositions comprising such recombinant DNA molecules, vectors or cells. Furthermore, the present invention relates to cells and transgenic non-human animals, comprising the aforementioned recombinant DNA molecules or vectors stably integrated into their genome and their use for the identification of substances capable of suppressing or activating transcription of a gene in endothelial cells. The present invention also relates to the use of the before described recombinant DNA molecules and vectors for the preparation of pharmaceutical compositions for treating, preventing, and/or delaying a vascular or tumorous disease in a subject. Furthermore, the recombinant DNA molecules and vectors of the invention can be used for the preparation of pharmaceutical compositions for inducing a vascular or tumorous disease in a non-human animal.

In the field of neuroscience and medical therapy, there is a great demand for test systems to study the function and interaction of gene products, the malfunction or expression of which cause vascular and/or tumorous diseases. Such systems would

also be suitable for drug development against such diseases. A prominent example for gene products involved in vascular diseases are angiogenic growth factors and their endothelial receptors which play a major role in the formation of the embryonic vascular system and in certain angiogenesis-dependent diseases, such as solid tumor growth or retinopathy. The Kinase-insert Domain-containing Receptor/fetal liver kinase-1 (KDR/Flk-1) in the following referred to as Flk-1 and Flt-1 are high affinity signaling receptors for the endothelial mitogen, vascular endothelial growth factor (VEGF) (Connolly, J. Clin. Invest. 84 (1989), 1470-1478; Leung, Science 246 (1989), 1306-1309). Through interactions with its receptors, VEGF plays critical roles in growth and maintenance of vascular endothelial cells and in the development of new blood vessels in physiologic and pathologic states (Aiello, New Engl. J. Med. 331 (1994), 1480-1487; Shweiki, Nature 359 (1992), 843-845; Berkman, J. Clin. Invest. 91 (1993), 153-159). The patterns of embryonic expression of VEGF suggest that it is crucial for differentiation of endothelial cells from hemangioblasts and for development of blood vessels at all stages of growth (Jakeman, Endocrinology 133 (1993), 848-859; Breier, Development 114 (1992), 521-532). Among many potentially angiogenic factors, VEGF is the only one with patterns of expression, secretion, and activity that suggest a specific angiogenic function in normal development (Klagsbrun, Current Biology 3 (1993), 699-702).

High-affinity receptors for VEGF are found only on endothelial cells, and VEGF binding has been demonstrated on macro- and microvascular endothelial cells and in quiescent and proliferating endothelial cells (Jakeman, Endocrinology 133 (1993), 848-859; Jakeman, Clin. Invest. 89 (1992), 244-253). The Flk-1 and Flt-1 have been identified as candidate VEGF receptors by affinity cross-linking and competition-binding assays (de Vries, Science 255 (1992), 989-991; Millauer, Cell 72 (1993), 835-846; Terman, Biochem. Biophys. Res. Commun. 187 (1992), 1579-1586). These two receptor tyrosine kinases contain seven similar extracellular immunoglobulin domains and a conserved intracellular tyrosine kinase domain interrupted by a kinase insert (de Vries, Science 255 (1992), 989-991; Matthews, Proc. Natl. Acad. Sci. U.S.A 88 (1991), 9026-9030; Terman, Oncogene 6 (1991), 1677-1683); they are expressed specifically by endothelial cells *in vivo* (Millauer,



Cell 72 (1993), 835-846; Peters, Proc. Natl. Acad. Sci. USA 90 (1993), 7533-7537; Yamaguchi, Development 118 (1993), 489-498). *In situ* hybridization in the developing mouse has demonstrated that Flk-1 is expressed in endothelial cells at all stages of development, as well as in the blood island in which endothelial cell precursors first appear (Millauer, Cell 72 (1993), 835-846). Flk-1 is a marker for endothelial cell precursors at their earliest stages of development (Yamaguchi, Development 118 (1993), 489-498).

The vascular endothelium is critical for physiologic responses including thrombosis and thrombolysis, lymphocyte and macrophage homing, modulation of the immune response, and regulation of vascular tone. The endothelium is also intimately involved in the pathogenesis of vascular diseases such as atherosclerosis (Ross, Nature 362 (1993), 801-809). Although a number of genes expressed in the endothelium have been characterized (Collins, J. Biol. Chem. 266 (1991), 2466-2473; Iademaro, J. Biol. Chem. 267 (1992), 16323-16329; Jahroudi, Mol. Cell. Biol. 14 (1994), 999-1008; Lee, J. Biol. Chem. 265 (1990), 10446-10450), expression of these genes is either not limited to vascular endothelium (e.g., the genes encoding von Willebrand factor, endothelin-1, vascular cell adhesion molecule-1), or is restricted to specific subpopulations of endothelial cells (e.g., the gene for endothelial-leukocyte adhesion molecule-1). Flk-1 (also known as VEGF-receptor 2) is expressed in endothelial cells during embryonic and postnatal development. The Flk-1 receptor is the first endothelial receptor to be expressed in endothelial cell precursors during embryonic vascular development. Gene targeting experiments in transgenic mice have demonstrated that this receptor is essential for endothelial cell differentiation (Shalaby, Nature 376 (1995), 62-66). Furthermore, in a variety of tumors, Flk-1 receptor expression is re-induced in the tumor vasculature, and it has been shown that signaling via the Flk-1 receptor is required for tumor vascularization and growth (Millauer, Nature 367 (1994), 576-579).

Recently, *in vitro* studies with the upstream region of the human Flk-1 gene (Patterson, J. Biol. Chem. (1995), 23111-23118) showed that DNA fragments located in the 5' flanking region of the human Flk-1 gene mediate expression of a reporter gene.

For studying all aspects of genes involved in vascular diseases such as atherosclerosis, however, the system described by Patterson, supra, suffers from several drawbacks. For example, promoter activity of the 5'-flanking region used by Patterson was also observed in cell types which do not express the Flk-1 gene naturally. Furthermore, the promoter fragment employed in Patterson, supra, was not shown to be expressed in vivo in its natural background. In order to specifically suppress or confer endothelium specific gene expression and for the development of endothelium specific drugs, however, one needs test systems which closely resemble the regulation of the Flk-1 expression in vivo since otherwise non-informative or even false positive results may be obtained.

Thus, the technical problem of the present invention is to provide means and methods that allow the modulation of gene expression specifically in endothelial cells in vivo, preferably at all stages of development.

The solution to this technical problem is achieved by providing the embodiments characterized in the claims.

Accordingly, the invention relates to a recombinant DNA molecule comprising:

- (a) a first regulatory sequence of an intron of the Vascular Endothelial Growth factor (VEGF) receptor-2 (Flk-1) gene or of a gene homologous to the Flk-1 gene being capable of conferring expression in endothelial cells in vivo; and
- (b) operatively linked thereto a heterologous DNA sequence.

In accordance with the present invention, a regulatory sequence driving the expression of a heterologous DNA sequence in substantially all endothelial cells in vivo, preferably at substantially all stages of development has been identified. Said regulatory sequence is suitable to direct the expression of a heterologous DNA sequence in the above-mentioned cells. The recombinant DNA molecule of the invention allows studying the function and interaction of proteins which are expressed in the endothelium of, for example, humans and the malfunction, and/or

unregulated expression of which is supposed to be the or a causative agent of vascular and tumorous diseases. Thus, the regulatory sequences of the invention are particularly suited and useful for the engineering of transgenic cells and non-human animals which can serve as a test system for the development of drugs for the treatment of vascular and tumorous diseases of endothelial origin.

In the context of the present invention the term "a first regulatory sequence of an intron of the Flk-1 gene" means a nucleotide sequence of the first intron of the murine Flk-1 gene including the regulatory sequences which are capable of conferring the specific expression of a heterologous DNA sequence in endothelial cells, preferably at all stages of development.

The high affinity receptor for vascular endothelial growth factor (VEGF), Flk-1, is the first endothelial receptor to be expressed in angioblast precursors, and its function is essential for the differentiation of the hemangioblastic lineage. In accordance with the present invention cis-acting regulatory elements of the murine Flk-1 gene have been identified that mediate endothelial specific expression of a reporter gene in transgenic mice. Sequences within the 5'-flanking region of the Flk-1 gene, in combination with sequences located in the first intron, specifically and reproducibly targeted transgene expression to endothelial cells of the embryonic vasculature. These sequences were capable of targeting expression of the heterologous DNA sequence to angioblasts during early stages of vascular development and also to the vasculature of postnatal mice. The regulatory sequences located in the first intron also functioned as an autonomous endothelium-specific enhancer when fused to the heterologous herpes simplex virus-thymidine kinase promoter. This Flk-1 intron enhancer contains several potential binding sites for transcription factors of the Ets and GATA families. Sequences of the Flk-1 promoter contributed to a strong, complete and reproducible endothelial cell-specific gene expression in the embryo and are essential for expression in the yolk sac.

In order to characterize cis-acting regulatory sequences contained in the Flk-1 gene, recombinant bacteriophage lambda clones containing mouse genomic DNA (Mouse strain 129/Sv) have been isolated encompassing a 21 kilo base pair (kb) region of

the mouse Flk-1 gene, contained in the DNA insertions of  $\lambda$  phages 6 and 16, including approximately 15 kb of 5' flanking sequences, exons 1, 2 and 3, and introns 1 and 2 (Fig. 4A). The DNA sequence of a 12.8 kb region spanning from position -6.65 kb (the affixes - and + refer to the nucleotide position relative to the transcriptional start site as shown in Fig. 1 which corresponds to nucleotide position 6661 of SEQ ID NO: 1) to position +6.15 (located in the third exon) was determined (SEQ ID NO: 1). Reporter gene studies were performed in order to characterize regulatory cis-acting elements of the Flk-1 gene. Initial studies focused on the role of 5' flanking sequences of the Flk-1 gene ("Flk-1 promoter") in mediating endothelium-specific expression in cultured bovine aortic endothelial (BAE) cells (Röncke, Circulation Research 79 (1996), 277-285). In these studies, it was found that a promoter fragment ranging from -624 to +299 mediated high expression of the luciferase reporter gene following transient transfection in BAE cells. Experiments with transgenic mouse embryos performed in accordance with the present invention revealed, however, that the murine promoter DNA fragments were not sufficient to mediate endothelium-specific reporter gene expression in vivo. Surprisingly, however, when a Flk-1 promoter fragment (ranging from, e.g., -624 to +299 bp) was combined with a 2.3 kb fragment of the first Flk-1 intron, endothelium specific expression of a lacZ reporter gene in mouse embryos was obtained. Thus, the first intron (nucleotides 7027 to 10642 of SEQ ID NO: 1) of the mouse Flk-1 gene is essential for endothelium specific gene expression. In particular, a DNA fragment (see Figure 12) comprising nucleotides 10094 to 10608 of SEQ ID NO: 1 was shown in accordance with the present invention to be sufficient to direct the expression of a heterologous DNA sequence into endothelial cells; see Example 8. This is a novel finding because the sequences described in previous publications (Patterson, supra; Röncke, supra) are, in contrast to the expectations and interpretations of the prior art, not sufficient to mediate endothelium-specific expression in vivo. These results obtained in accordance with the present invention demonstrate that the regulatory sequences located in the intron of the Flk-1 gene mediating endothelium-specific expression can be used to direct expression of heterologous genes in the vasculature.

The genomic DNA of the murine Flk-1 gene comprising the intron regulatory sequences can be obtained from liver of mouse strain 129/SV, or, for example, by screening a phage library of liver genomic DNA in the vector  $\lambda$ FixII (Stratagene, La Jolla, CA) generated by conventional methods known in the art.

The term "regulatory sequence of a gene homologous to the Flk-1 gene" also includes promoter regions and regulatory sequences of a gene from another species, for example, humans and other mammals which is homologous to the Flk-1 gene of mouse and which confers the same or substantially the same expression pattern. Such regulatory sequences are characterized by their capability of conferring expression of a heterologous DNA sequence specifically in endothelial cells in vivo, preferably at all stages of development. Thus, according to the present invention, regulatory sequences from other species can be used that are functionally homologous to the regulatory sequences of the intron of the Flk-1 gene from mouse, or regulatory sequences of genes that display a substantially identical pattern of expression, in the sense of being expressed in the endothelium, preferably at all stages of development.

It is possible for the person skilled in the art to isolate by employing the known Flk-1 gene from mouse, corresponding genes from other species, for example, humans and other mammals. This can be done by conventional techniques known in the art, for example, by using Flk-1 gene sequences as a hybridization probe or by designing appropriate PCR primers. It is then possible to isolate the corresponding regulatory sequences by conventional techniques and test them for their expression pattern. For this purpose, it is, for instance, possible to fuse the regulatory sequences to a reporter gene, such as the luciferase or green fluorescent protein (GFP) encoding genes and assess the expression of the reporter gene in transgenic animals, for example in mice. The partial nucleotide sequence of the human Flk-1 gene may be obtained from Genbank Acc. No. X89776; Patterson, supra; Terman, Biochem. Biophys. Res. Comm. 187 (1992), 1579-1586; Genbank Acc. No. X61656. The present invention also relates to recombinant DNA molecules comprising regulatory sequences which are substantially identical to that of the Flk-1 intron or to

an intron of a homologous gene or to fragments thereof and which are able to confer specific expression in endothelial cells, preferably at all stages of development in mouse or other mammals.

Such regulatory sequences differ at one or more positions from the above-mentioned regulatory sequences but still have the same specificity, namely they comprise the same or similar sequence motifs responsible for the above described expression pattern. Preferably such regulatory sequences hybridize to one of the above-mentioned regulatory sequences, most preferably under stringent conditions. Particularly preferred are regulatory sequences which share at least 85%, more preferably 90-95%, and most preferably 96-99% sequence identity with one of the above-mentioned regulatory sequences and have the same specificity. Such regulatory sequences also comprise those which are analogues or derivatives, for example by way of nucleotide deletion(s), insertion(s), substitution(s), addition(s), and/or recombination(s) and/or any other modification(s) known in the art either alone or in combination in comparison to the above-described nucleotide sequence. Methods for introducing such modifications in the nucleotide sequence of the regulatory sequences of the invention are well known to the person skilled in the art and described, for example, in Sambrook (Molecular cloning; A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY (1989)). All such fragments, analogues and derivatives of the regulatory sequence of the invention are included within the scope of the present invention, as long as the essential characteristic regulatory properties as defined above remain unaffected in kind. It is also immediately evident to the person skilled in the art that further regulatory sequences may be added to the regulatory sequences of the invention. For example promoters, transcriptional enhancers and/or sequences which allow for induced expression of the regulatory sequences of the invention may be employed. A suitable inducible system is for example tetracycline-regulated gene expression which is described by, e.g., Gossen (Proc. Natl. Acad. Sci. USA 89 (1992), 5547-5551; Trends Biotech. 12 (1994), 58-62).

The expression conferred by the regulatory sequences of the invention may not be exclusively limited to the above-described specificity but may also occur in, e.g.,

neuronal cells, including neural retinal progenitor cells at all or different stages of development and haematopoietic cells (Yang, J. Neurosci. 16 (1996), 6089-6099).

The term "further regulatory sequences" refers to sequences which influence the specificity and/or level of expression, for example in the sense that they confer cell and/or tissue specificity or developmentally and/or inducible regulated gene expression. Such regions can be located upstream of or comprising the transcription initiation site, such as a promoter, but can also be located downstream of it, e.g., in transcribed but not translated leader sequences.

The term "promoter" refers to the nucleotide sequences necessary for transcription initiation, i.e. RNA polymerase binding, and also includes, for example, the TATA box.

The term "in vivo" for the purpose of the present invention is used for cells in an organism as opposed to cells growing in culture (in vitro).

The term "heterologous" with respect to the DNA sequence being operatively linked to the promoter of the invention means that said DNA sequence is not naturally linked to the regulatory sequences comprised in the recombinant DNA molecule of the invention.

In a preferred embodiment said first regulatory sequence of the invention comprises a GATA-binding site, an AP-1 binding site, an SP1 binding site, site, an NFκB binding site, a STAT binding site, a Scl/Tal-1 binding site, an Ets-1 binding site, a PEA3 consensus sequence or any combination(s) thereof. A functional analysis of the first 6.5 kbp of the transcribed region of the murine Flk-1 genes lead to the identification of an endothelial-specific positive regulatory element. This regulatory sequence is located in the region between the XhoI and BamHI restriction site in the first intron of the Flk-1 gene (cf. Fig. 4A). It is functional in both orientations since the intron enhancer was used in an antiparallel manner with respect to the Flk-1 promoter fragment in the construct referred to as 3'-In 1; see Example 2 hereinbelow. A sequence analysis of the intron lead to the identification of two potential GATA binding sites (+1927 Bp, +3514 Bp); a potential AP-1 binding site

(+2210 Bp) and two PEA3 consensus sequences (+3494 Bp, +3741 Bp); see Fig. 1. As demonstrated in Example 8, the intron sequences that were sufficient for endothelium-specific expression were contained in a 510 bp fragment (nucleotides 10094 to 10608 of SEQ ID NO: 1). Several potential binding sites for known transcription factors could be identified therein (see Figure 12), including consensus binding sites for c-ets1, PEA3 (an Ets-like transcription factor), GATA transcription factors, and Scl/Tal-1. The c-ets1 transcription factor was proposed to be involved in the early differentiation of endothelial cells from their precursors (Pardanaud Cell Adhesion and Communication 1 (1993), 151-160). In addition, c-ets1 is expressed in endothelial cells during tumor vascularization and other forms of angiogenesis in humans (Wernert, Am. J. Pathol. 140 (1992), 119-127). Proteins of the Ets family can activate transcription through a PEA3 motif (Wernert, 1992). Transcription factors of the GATA family are involved in the transcription of genes that are expressed in the hematopoietic and endothelial lineages, such as *von Willebrand factor* (Jahroudi, Mol. Cell. Biol. 14 (1994), 999-1008). Unlike the hematopoietic-transcription factor GATA-1, GATA-2 is expressed in both the endothelial and hematopoietic lineages (Elefanti, Blood 90 (1997), 1435-1447). Scl/Tal-1 has recently been implicated in the regulation of *Flk-1* expression in Zebrafish (Liao, Genes Dev. 12 (1998), 621-626). The presence of two potential Scl/Tal-1 binding sites in the murine *Flk-1* intron enhancer suggests that Scl/Tal-1 might regulate *Flk-1* expression in mice. However, no direct effect of Scl/Tal-1 on *Flk-1* expression has been observed so far in mice, although Scl-null mice have vascular defects (Visvader, Genes Dev. 12 (1998), 473-479). Knock out experiments performed with the above-described regulatory sequences will easily reveal which of these elements present in, e.g., the 510 bp fragment (nucleotides 10094 to 10608 of SEQ ID NO: 1) are involved in the control of the regulatory sequence and the sequential order of these elements necessary to confer endothelium specific gene expression. Of course, the regulatory sequences obtained from such studies are also within the scope of the present invention.

Preferably, said first regulatory sequence is selected from the group consisting of



- (a) DNA sequences comprising a nucleotide sequence as given in SEQ ID NO: 1;
- (b) DNA sequences comprising a nucleotide sequence of SEQ ID NO: 1 from nucleotide 8260 to nucleotide 10560, from nucleotide 8336 to nucleotide 10608 and/or from nucleotide 10094 to nucleotide 10608;
- (c) DNA sequences comprising the nucleotide sequence of the human Flk-1-intron;
- (d) DNA sequences comprising a nucleotide sequence which hybridizes with a nucleotide sequence of (a), (b) or (c) under stringent conditions;
- (e) DNA sequences comprising a nucleotide sequence which is conserved in the nucleotide sequences of (a), (b) and (c); and
- (f) DNA sequences comprising a fragment, analogue or derivative of a nucleotide sequence of any one of (a) to (e) capable of conferring expression in endothelial cells.

In a particularly preferred embodiment of the present invention, the regulatory sequences comprise the nucleotides 8260 to 10560, 8336 to 10608 (comprising the BamHI/XhoI fragment of the first intron (+1677 bp/+3947); see Figure 4 and Examples 1 to 10), most preferably nucleotides 8560 to 10400 and still more preferably nucleotides 10094 to 10608 (comprising the Swal/BamHI fragment (+3437 bp/3947 bp); see Example 8) of the nucleotide sequence as set forth in SEQ ID No. 1 or a fragment thereof, which still confers expression in endothelial cells, preferably at all stages of development.

In a further preferred embodiment of the invention the heterologous DNA sequence of the recombinant DNA molecules described above is operatively linked to further regulatory sequences. Expression comprises transcription of the nucleic acid molecule, preferably into a translatable mRNA. Regulatory elements ensuring expression in eukaryotic cells, preferably mammalian cells, are well known to those skilled in the art. They normally comprise promoters ensuring initiation of transcription and optionally poly-A signals ensuring termination of transcription and stabilization of the transcript. Additional regulatory elements may include

transcriptional as well as translational enhancers. Preferably said further regulatory sequence is a promoter and/or a 3'-untranslated region.

Although some endothelial-specific promoters have been characterized, e.g. of the genes for von Willebrand factor (Jahroudi, Mol. Cell Biol. 14 (1994), 999-1008), Endothelin-1 (Lee, J. Biol. Chem. 265 (1990), 10446-10450), E-selectin (Collins, J. Biol. Chem. 266 (1991), 2466-2473), Tie-2 (Schlaeger, Development 121 (1995), 1089-1098), VCAM-1 (Iademarco, J. Biol. Chem. 267 (1992), 16323-16329) and endothelial NO-synthase (Zhang, J. Biol. Chem. 270 (1995), 15320-15326) these genes are neither specific for proliferating endothelium, nor necessary for endothelial cell determination. Due to the present invention these promoters can now be combined with the regulatory sequences of the invention in order to mediate endothelium specific gene expression of heterologous DNA sequences. However, other promoters can be used as well. For example, it is shown in Example 8 that the regulatory sequences of the invention conferred endothelium-specific gene expression to the heterologous herpes simplex virus-thymidine kinase (tk) promoter.

In a preferred embodiment the above mentioned promoter is a promoter of hypoxia inducible genes, genes encoding growth factors such as VEGF, PDGF or Fibroblast growth factor or their receptors or glycolytic enzymes.

In a particularly preferred embodiment said promoter comprises a DNA sequence selected from the group consisting of

- (a) DNA sequences comprising the nucleotide sequence as given in SEQ ID NO:1 from nucleotide 6036 to nucleotide 6959;
- (b) DNA sequences comprising the nucleotide sequence of the human Flk-1 promoter;
- (c) DNA sequences comprising a nucleotide sequence which hybridizes with a nucleotide sequence of (a) or (b) under stringent conditions;
- (d) DNA sequences comprising a nucleotide sequence which is conserved in the nucleotide sequences of (a) and (b); and

- (e) DNA sequences comprising a fragment, analogue or derivative of a nucleotide sequence of any one of (a) to (d).

At least one of the aforescribed DNA sequences may be preferably of human or murine origin although other sources may be employed as well. Preferably, the heterologous DNA sequence being operatively linked to the regulatory sequences is located 5' to the regulatory sequence of the invention.

In a further preferred embodiment, the heterologous DNA sequence of the above-described recombinant DNA molecules encodes a peptide, protein, antisense RNA, sense RNA and/or ribozyme. The recombinant DNA molecule or vector of the invention can be used alone or as part of a vector to express heterologous DNA sequences, which, e.g., encode proteins other than Flk-1, in cells of the blood vessel wall, i.e., endothelial cells, for, e.g., gene therapy or diagnostics of vascular diseases such as atherosclerosis. The recombinant DNA molecule or vector containing DNA sequence encoding a protein of interest is introduced into endothelial cells which in turn produce the protein of interest. For example, sequences encoding t-PA (Pennica, Nature 301 (1982), 214), p21 cell cycle inhibitor (El-Deiry, Cell 75 (1993), 817-823), or nitric oxide synthase (Bredt, Nature 347 (1990), 768-770) may be operatively linked to the endothelial cell-specific regulatory sequences of the invention and expressed in endothelial cells. For example, thrombolytic agents can be expressed under the control of the endothelial cell-specific regulatory sequences of the invention for expression by vascular endothelial cells in blood vessels, e.g., vessels occluded by aberrant blood clots. Other heterologous proteins, e.g., proteins which inhibit smooth muscle cell proliferation, e.g., interferon- $\gamma$  and atrial natriuretic polypeptide, may be specifically expressed in endothelial cells to ensure the delivery of these therapeutic peptides to an atherosclerotic lesion or an area at risk of developing an atherosclerotic lesion, e.g., an injured blood vessel.

The endothelial cell-specific regulatory sequences of the invention may also be used in gene therapy to promote angiogenesis to treat diseases such as peripheral vascular disease or coronary artery disease (Isner, Circulation 91 (1995), 2687-

2692). For example, the regulatory sequences of the invention can be operatively linked to sequences encoding cellular growth factors which promote angiogenesis, e.g., VEGF, acidic fibroblast growth factor, basic fibroblast growth factor and the like.

In a most preferred embodiment of the present invention, said protein is selected from the group consisting of Vascular Endothelial Growth Factor (VEGF), Hypoxia Inducible Factors (HIF), HIF-Related Factor (HRF), tissue plasminogen activator, p21 cell cycle inhibitor, nitric oxide synthase, interferon- $\gamma$ , atrial natriuretic polypeptide and monocyte chemotactic proteins.

In another particularly preferred embodiment of the invention, said protein is a scorable marker, preferably luciferase, green fluorescent protein or lacZ. This embodiment is particularly useful for simple and rapid screening methods for compounds and substances described herein below capable of modulating the expression of genes in the endothelium. For example, endothelial cells can be cultured with VEGF in the presence and absence of the candidate compound in order to determine whether the compound affects the expression of genes which are under the control of regulatory sequences of the invention, which can be measured, e.g., by monitoring the expression of the above-mentioned marker. It is also immediately evident to those skilled in the art that other marker genes may be employed as well, encoding, for example, selectable marker which provide for the direct selection of compounds which induce or inhibit the expression of said marker.

The regulatory sequences of the invention may also be used in methods of antisense therapy. Antisense therapy may be carried out by administering to an animal or a human patient, a recombinant DNA containing the endothelial cell-specific regulatory sequences of the invention operably linked to a DNA sequence, i.e., an antisense template which is transcribed into an antisense RNA. The antisense RNA may be a short (generally at least 10, preferably at least 14 nucleotides, and up to 100 or more nucleotides) nucleotide sequence formulated to be complementary to a portion of a specific mRNA sequence. Standard methods

relating to antisense technology have been described (Melani, Cancer Res. 51 (1991), 2897-2901). Following transcription of the DNA sequence into antisense RNA, the antisense RNA binds to its target mRNA molecules within a cell, thereby inhibiting translation of the mRNA and down-regulating expression of the protein encoded by the mRNA. For example, an antisense sequence complementary to a portion of or all of the Flk-1 (KDR) mRNA (Terman, Oncogene 6 (1991), 1677-1683 and Terman (1992), supra) would inhibit the expression of Flk-1, which in turn would inhibit angiogenesis. Such antisense therapy may be used to treat cancer, particularly to inhibit angiogenesis at the site of a solid tumor, as well as other pathogenic conditions which are caused by or exacerbated by angiogenesis, e.g., inflammatory diseases such as rheumatoid arthritis, and diabetic retinopathy.

The expression of other endothelial cell proteins may also be inhibited in a similar manner, for example, endothelial cell proteins such as cell cycle proteins (thereby inhibiting endothelial cell proliferation, and therefore, angiogenesis); coagulation factors such as von Willebrand factor; and endothelial cell adhesion factors, such as ICAM-1 and VCAM-1 (Bennett, J. Immunol. 152 (1994), 3530-3540).

Thus, in a further preferred embodiment of the present invention, said antisense RNA or said ribozyme are directed against a gene involved in vasculogenesis and/or angiogenesis and/or tumors of endothelial origin.

In a further embodiment, the invention relates to nucleic acid molecules of at least 15 nucleotides in length hybridizing specifically with a regulatory sequence as described above or with a complementary strand thereof. This means that they hybridize, preferably under stringent conditions, specifically with the nucleotide sequences as described above and show no or very little cross-hybridization with nucleotide sequences having no or substantially different regulatory properties. Such nucleic acid molecules may be used as probes and/or for the control of gene expression. Nucleic acid probe technology is well known to those skilled in the art who will readily appreciate that such probes may vary in length. Preferred are nucleic acid probes of 17, 18, 19, 20 to 25 and 25 to 35 nucleotides in length. Of course, it may also be appropriate to use nucleic acids of up to 100 and more

nucleotides in length. The nucleic acid probes of the invention are useful for various applications. On the one hand, they may be used as PCR primers for amplification of regulatory sequences according to the invention. Another application is the use as a hybridization probe to identify regulatory sequences hybridizing to the regulatory sequences of the invention by homology screening of genomic DNA libraries. Nucleic acid molecules according to this preferred embodiment of the invention which are complementary to a regulatory sequence as described above may also be used for repression of expression of a gene comprising such regulatory sequences, for example due to an antisense or triple helix effect or for the construction of appropriate ribozymes (see, e.g., EP-B1 0 291 533, EP-A1 0 321 201, EP-A2 0 360 257) which specifically cleave the (pre)-mRNA of a gene comprising a regulatory sequence of the invention. Selection of appropriate target sites and corresponding ribozymes can be done as described for example in Steinecke, Ribozymes, Methods in Cell Biology 50, Galbraith et al. eds Academic Press, Inc. (1995), 449-460. Furthermore, the person skilled in the art is well aware that it is also possible to label such a nucleic acid probe with an appropriate marker for specific applications, such as for the detection of the presence of a nucleic acid molecule of the invention in a sample derived from an organism.

Such molecules may either be DNA or RNA or a hybrid thereof. Furthermore, said nucleic acid molecule may contain, for example, thioester bonds and/or nucleotides analogues, commonly used in oligonucleotide anti-sense approaches. Said modifications may be useful for the stabilization of the nucleic acid molecule against endo- and/or exonucleases in the cell. Said nucleic acid molecules may also be transcribed by an appropriate vector containing a chimeric gene which allows for the transcription of said nucleic acid molecule in the cell. Such nucleic acid molecules may further contain ribozyme sequences which specifically cleave the (pre)-mRNA comprising the regulatory sequence of the invention. Furthermore, oligonucleotides can be designed which are complementary to a regulatory sequence of the invention (triple helix; see Lee, Nucl. Acids Res. 6 (1979), 3073; Cooney, Science 241 (1988), 456 and Dervan, Science 251 (1991), 1360), thereby preventing transcription and the production of the encoded protein.

The present invention also relates to vectors, particularly plasmids, cosmids, viruses, bacteriophages used conventionally in genetic engineering that comprise a recombinant DNA molecule of the invention. Preferably, said vector is an expression vector and/or a targeting vector. Expression vectors derived from viruses such as retroviruses, vaccinia virus, adeno-associated virus, herpes viruses, or bovine papilloma virus, may be used for delivery of the recombinant DNA molecule or vector of the invention into targeted cell population. Methods which are well known to those skilled in the art can be used to construct recombinant viral vectors; see, for example, the techniques described in Sambrook, Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory (1989) N.Y. and Ausubel, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y. (1989). Alternatively, the recombinant DNA molecules and vectors of the invention can be reconstituted into liposomes for delivery to target cells.

As is demonstrated in Example 11, the Flk-1 promoter was stimulated by HIF-2 $\alpha$ , a basic helix-loop-helix/PAS domain transcription factor related to hypoxia-inducible factor-1. HIF-2 $\alpha$  has previously been shown to stimulate the expression of VEGF, suggesting that HIF-2 $\alpha$  may regulate the coordinate expression of both the VEGF receptor Flk-1 and its ligand in vivo. Thus, Flk-1 gene regulatory elements described herein can be used together with HIF-2 $\alpha$  for the elucidation of the molecular mechanisms involved in endothelial cell specification and angiogenesis, and can be used to target expression of any transgene to the endothelium. Thus, in a preferred embodiment, the vector of the invention further comprises a gene capable of expressing HIF-2 $\alpha$ .

The present invention furthermore relates to host cells transformed with a DNA molecule or vector of the invention. Said host cell may be a prokaryotic or eukaryotic cell. The vector or recombinant DNA molecule of the invention which is present in the host cell may either be integrated into the genome of the host cell or it may be maintained extrachromosomally. In this respect, it is also to be understood that the recombinant DNA molecule of the invention can be used for "gene targeting" and/or

"gene replacement", for restoring a mutant gene or for creating a mutant gene via homologous recombination.

The host cell can be any prokaryotic or eukaryotic cell, such as a bacterial, insect, fungal, plant or animal cell. Preferred fungal cells are, for example, those of the genus *Saccharomyces*, in particular those of the species *S. cerevisiae*. Suitable mammalian cell lines comprise Saos-2 human osteosarcoma cells (ATCC HTB-85), HeLa human epidermoid carcinoma cells (ATCC CRL-7923), HepG2 human hepatoma cells (ATCC HB-8065), human fibroblasts (ATCC CRL-1634), U937 human histiocytic lymphoma cells (ATCC CRL-7939), RD human embryonal rhabdomyosarcoma cells (ATCC CCL-136), MCF7 human breast adenocarcinoma cells (ATCC HTB-22), JEG-3 human choriocarcinoma cells (ATCC HB36), A7r5 fetal rat aortic smooth muscle cells (ATCC CRL-1444), and NIH 3T3 mouse fibroblasts (ATCC CRL-1658) obtainable from the American Type Culture Collection. Primary-culture HUVEC may be obtained from Clonetics Corp. (San Diego, CA) and can be grown in EGM medium containing 2% fetal calf serum (Clonetics). Primary-culture human aortic and intestinal smooth muscle cells can also be obtained from Clonetics Corp. Most preferably said host cell is an endothelial cell or derived therefrom, such as BAE cells. In view of the synergistic effect of the co-expression of a recombinant DNA molecule of the invention and HIF-2 $\alpha$ , a further embodiment of the invention concerns the above-described cells which further comprise a recombinant DNA molecule or vector containing a gene capable of expressing HIF-2 $\alpha$ .

Moreover, the present invention relates to a pharmaceutical composition comprising at least one of the aforementioned recombinant DNA molecules or vectors of the invention, either alone or in combination, and optionally a pharmaceutically acceptable carrier or excipient. Examples of suitable pharmaceutical carriers are well known in the art and include phosphate buffered saline solutions, water, emulsions, such as oil/water emulsions, various types of wetting agents, sterile solutions etc. Compositions comprising such carriers can be formulated by well known conventional methods. These pharmaceutical compositions can be administered to the subject at a suitable dose. Administration of the suitable compositions may be



effected by different ways, e.g. by intravenous, intraperitoneal, subcutaneous, intramuscular, topical or intradermal administration. The dosage regimen will be determined by the attending physician and other clinical factors. As is well known in the medical arts, dosages for any one patient depends upon many factors, including the patient's size, body surface area, age, the particular compound to be administered, sex, time and route of administration, general health, and other drugs being administered concurrently. Dosages will vary but a preferred dosage for intravenous administration of DNA is from approximately  $10^6$  to  $10^{22}$  copies of the DNA molecule. The compositions of the invention may be administered locally or systemically. Administration will generally be parenterally, e.g., intravenously; DNA may also be administered directly to the target site, e.g., by biolistic delivery to an internal or external target site or by catheter to a site in an artery.

It is envisaged by the present invention that the various recombinant DNA molecules and vectors of the invention are administered either alone or in any combination using standard vectors and/or gene delivery systems, and optionally together with an appropriate compound, for example VEGF, and/or together with a pharmaceutically acceptable carrier or excipient. Subsequent to administration, said recombinant DNA molecules may be stably integrated into the genome of the mammal. On the other hand, viral vectors may be used which are specific for certain cells or tissues, preferably for the endothelium and persist in said cells. Suitable pharmaceutical carriers and excipients are well known in the art. The pharmaceutical compositions prepared according to the invention can be used for the prevention or treatment or delaying of different kinds of diseases, which are related to the expression or overexpression of a given gene or genes in the endothelium.

Furthermore, it is possible to use a pharmaceutical composition of the invention which comprises a recombinant DNA molecule or vector of the invention in gene therapy. Suitable gene delivery systems may include liposomes, receptor-mediated delivery systems, naked DNA, and viral vectors such as herpes viruses, retroviruses, adenoviruses, and adeno-associated viruses, among others. Delivery of nucleic

acids to a specific site in the body for gene therapy or antisense therapy may also be accomplished using a biolistic delivery system, such as that described by Williams (Proc. Natl. Acad. Sci. USA 88 (1991), 2726-2729).

Standard methods for transfecting cells with recombinant DNA are well known to those skilled in the art of molecular biology, see, e.g., WO 94/29469. Gene therapy and antisense therapy to prevent or decrease the development of atherosclerosis or inhibit angiogenesis may be carried out by directly administering the recombinant DNA molecule or vector of the invention to a patient or by transfecting endothelial cells with the recombinant DNA molecule or vector of the invention *ex vivo* and infusing the transfected cells into the patient. Furthermore, research pertaining to gene transfer into cells of the germ line is one of the fastest growing fields in reproductive biology. Gene therapy, which is based on introducing therapeutic genes into cells by ex-vivo or in-vivo techniques is one of the most important applications of gene transfer. Suitable vectors and methods for in-vitro or in-vivo gene therapy are described in the literature and are known to the person skilled in the art; see, e.g., WO94/29469, WO 97/00957 or Schaper (Current Opinion in Biotechnology 7 (1996), 635-640) and references cited therein. The DNA molecules and vectors comprised in the pharmaceutical composition of the invention may be designed for direct introduction or for introduction via liposomes, or viral vectors (e.g. adenoviral, retroviral) containing said recombinant DNA molecule into the cell. Preferably, said cell is a germ line cell, embryonic cell, or egg cell or derived therefrom. The pharmaceutical compositions according to the invention can be used for the treatment of all kinds of diseases hitherto unknown as being related to the expression and/or over expression of genes in the endothelium.

The present invention also relates to diagnostic compositions or kits comprising at least one of the aforementioned recombinant DNA molecules, vectors, cells and/or nucleic acid molecules and, in the case of diagnostic compositions, optionally suitable means for detection.

Said diagnostic compositions may be used for methods of detecting and isolating regulatory sequences which are a functionally equivalent to the Flk-1 intron

regulatory sequences of the invention. The kits of the invention may further contain compounds such as further plasmids, antibiotics and the like for screening transgenic animals and/or animal cells useful for the genetic engineering of non-human animals, preferably mammals and most preferably mouse.

It is to be understood that the introduced recombinant DNA molecules and vectors of the invention express the heterologous DNA sequence after introduction into said cell and preferably remain in this status during the lifetime of said cell. For example, cell lines which stably express the heterologous DNA under the control of the regulatory sequence of the invention may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with the recombinant DNA molecule or vector of the invention and a selectable marker, either on the same or separate vectors. Following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the heterologous DNA sequence under the control of the regulatory sequence of the invention, and which respond to VEGF and/or hypoxia mediated signal transduction. Such engineered cell lines are particularly useful in screening compounds capable of modulating Flk-1 gene expression.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, Cell 11(1977), 223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska, Proc. Natl. Acad. Sci. USA 48 (1962), 2026), and adenine phosphoribosyltransferase (Lowy, Cell 22 (1980), 817) genes can be employed in tk<sup>-</sup>, hgprt<sup>-</sup> or aprt<sup>-</sup> cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, Proc. Natl. Acad. Sci. USA 77 (1980), 3567; O'Hare, Proc. Natl. Acad. Sci. USA 78 (1981), 1527), gpt, which confers resistance to mycophenolic acid (Mulligan, Proc. Natl. Acad. Sci. USA 78 (1981), 2072); neo, which confers resistance to the

aminoglycoside G-418 (Colberre-Garapin, J. Mol. Biol. 150 (1981), 1); and hygromycin (Santerre, Gene 30 (1984), 147) genes. Additional selectable genes have been described, namely *trpB*, which allows cells to utilize indole in place of tryptophan; *hisD*, which allows cells to utilize histinol in place of histidine (Hartman, Proc. Natl. Acad. Sci. USA 85 (1988), 8047); and ODC (ornithine decarboxylase) which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine, DFMO (McConlogue, 1987, In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed.). On the other hand, the person skilled in the art may also use the regulatory sequences of the invention to "knock out" an endogenous gene comprising identical or similar regulatory sequences, for example, by gene targeting, cosuppression, triple helix, antisense or ribozyme technology.

The present invention also relates to a method for the production of a transgenic animal, preferably mouse, comprising introduction of a recombinant DNA molecule or vector of the invention into a germ cell, an embryonic cell or an egg or a cell derived therefrom. The non-human animal to be used in the method of the invention may be a wildtype, i.e. healthy animal, or may have a disease or disorder, preferably a disease or disorder which is dependent on neovascularization, such as solid tumors, retinopathy, arthritis, psoriasis. Said disease or disorder may be an inborn insufficiency or naturally developed or caused by genetical engineering, for instance by the expression of a DNA sequence encoding a protein involved in neuronal development and/or diseases as described above, preferably under the control of the regulatory sequence of the invention.

The invention also relates to transgenic non-human animals comprising a recombinant DNA molecule or vector of the invention or obtained by the method described above, preferably wherein said recombinant DNA molecule is stably integrated into the genome of said non-human animal, preferably such that the presence of said recombinant DNA molecule or vector leads to the transcription and/or expression of the heterologous DNA sequence by the regulatory sequence of

the invention. Further non-human animals which may be employed according to the embodiments of the invention as described above are well known to the person skilled in the art and comprise rat, hamster, dog, monkey, rabbit, pig.

With the regulatory sequences of the invention, it is now possible to study *in vivo* the regulation of Flk-1 expression during angiogenesis. Furthermore, since VEGF and VEGF receptor genes seem to have different functions in different stages of development, it is now possible to determine domains of said proteins which may be important for their biological activity and/or for the regulation of their activity. In addition, it is now possible to in vivo study mutations which affect different functional or regulatory aspects of VEGF or its receptor or vector of the invention.

Moreover, the present invention relates to a method for the identification of a chemical and/or biological substance capable of suppressing or activating and/or enhancing the transcription of a gene in endothelial cells comprising:

- (a) contacting a cell of the invention or the transgenic non-human animal of the invention either of which is capable of expressing the heterologous DNA sequence with a plurality of compounds; and
- (b) determining those compounds which suppress or activate and/or enhance the expression of said heterologous DNA sequence.

Said plurality of compounds may be comprised in, for example, samples, e.g. cell extracts from, e.g. plants, animals or microorganisms. Furthermore, said compounds may be known in the art but hitherto not known to be capable of suppressing or activating and/or enhancing the transcription of a gene in endothelial cells. The plurality of compounds may be, e.g., added to the culture medium or injected into the animals.

The term "plurality of compounds" in a method of the invention is to be understood as a plurality of substances which are either identical or not. If a sample containing a plurality of compounds is identified in the method of the invention, then it is either possible to isolate the compound from the original sample identified as containing the compound capable of suppressing or activating and/or enhancing the

transcription of a gene in endothelial cells, or one can further subdivide the original sample, for example, if it consists of a plurality of different compounds, so as to reduce the number of different substances per sample and repeat the method with the subdivisions of the original sample. Depending on the complexity of the samples, this can be done several times, preferably until the sample identified according to the method of the invention only comprises a limited number of or only one substance(s). Preferably said sample comprises substances of similar chemical and/or physical properties, most preferably said substances are identical.

Determining whether a compound is capable of suppressing or activating and/or enhancing the transcription of a gene in endothelial cells can be done, for example, in mice by monitoring reporter gene expression or by monitoring behavior of the transgenic non-human animals of the invention contacted with the compounds compared to that of wild-type animals or compared to a transgenic non-human animal contacted with a compound which is either known to be capable or incapable of suppressing or activating and/or enhancing the transcription of a gene in endothelial cells of said transgenic non-human animal of the invention. Furthermore, the person skilled in the art can monitor the physical behavior, or for example the movement of the above-described animals. Such methods are well known in the art. Such regulators of Flk-1 gene expression may be used in processes such as wound healing; in contrast, antagonists of expression may be used in the treatment of tumors that rely on vascularization for growth. Thus, the present invention provides methods for identifying compounds which modulate VEGF receptor (e.g., Flk-1 or Flt1) gene expression. Compounds found to downregulate expression of a VEGF receptor gene can be used in methods to inhibit angiogenesis, while compounds found to enhance Flk-1 or Flt1 expression can be used in methods to promote angiogenesis, for example, to promote wound healing (e.g., healing of broken bones, burns, diabetic ulcers, and traumatic or surgical wounds) or to treat peripheral vascular disease, atherosclerosis, cerebral vascular disease, hypoxic tissue damage (e.g., retinopathy, hypoxic damage to heart tissue), diabetic pathologies such as chronic skin lesions, or coronary vascular disease. These compounds can also be used to treat patients who have, or have had, transient

ischemic attacks, vascular graft surgery, balloon angioplasty, frostbite, gangrene, or poor circulation. Compounds identified as having the desired effect (i.e., enhancing or inhibiting Flk-1 expression) can be tested further in appropriate models of endothelial cell growth and angiogenesis which are known to those skilled in the art. Given the therapeutic value of the compounds identified in accordance with the above-described method the present invention also relates to a method for the production of a pharmaceutical composition comprising the steps of the method of the invention and formulating the compound identified in step (b) in a pharmaceutically acceptable form.

The therapeutic compounds identified using the method of the invention may be administered to a patient by any appropriate method for the particular compound, e.g., orally, intravenously, parenterally, transdermally, transmucosally, or by surgery or implantation (e.g., with the compound being in the form of a solid or semi-solid biologically compatible and resorbable matrix) at or near the site where the effect of the compound is desired. For example, a salve or transdermal patch that can be directly applied to the skin so that a sufficient quantity of the compound is absorbed to increase vascularization locally may be used. This method would apply most generally to wounds on the skin. Salves containing the compound can be applied topically to induce new blood vessel formation locally, thereby improving oxygenation of the area and hastening wound healing. Therapeutic doses are determined to be appropriate by one skilled in the art.

Furthermore, identification of transacting factors which interact with the regulatory sequences of the invention can form the basis for the development of novel therapeutics for modulating conditions associated with endothelial cell growth, such as angiogenesis, vascular disease, and wound healing. Identification of transacting factors is carried out using standard methods in the art (see, e.g., Sambrook, supra, and Ausubel, supra). To determine whether a protein binds to the regulatory sequences of the invention standard DNA footprinting and/or native gel-shift analyses can be carried out. In order to identify the transacting factor which binds to the regulatory sequence of the invention, the regulatory sequences can be used as

an affinity reagent in standard protein purification methods, or as a probe for screening an expression library. Once the transacting factor is identified, modulation of its binding to the regulatory sequence in the Flk-1 gene can be pursued, beginning with, for example, screening for inhibitors of transacting factor binding. Enhancement of Flk-1 expression in a patient, and thus enhancement of angiogenesis, may be achieved by administration of the transacting factor, or the gene encoding it (e.g., in a vector for gene therapy). In addition, if the active form of the transacting factor is a dimer, dominant-negative mutants of the transacting factor could be made in order to inhibit its activity. Furthermore, upon identification of the transacting factor, further components in the pathway of Flk-1 signal transduction can be identified. Modulation of the activities of these components can then be pursued, in order to develop additional drugs and methods for modulating endothelial cell growth and angiogenesis.

As discussed in the background section of the description of the present invention, the interaction of VEGF and its receptor play an important role in the onset of angiogenic disease. Transgenic non-human animals expressing VEGF and/or its receptor gene and/or mutated versions thereof under the control of the regulatory sequences of the invention can now be used for the identification of substances, which, for example, are capable of restoring the wild-type interaction of mutated VEGF and its receptor either or both of which bear mutations. Some genetic changes lead to altered protein conformational states. Genetic changes may therefore result in a decreased binding activity of VEGF. Restoring the activity of mutant VEGF protein or increasing the activity of other proteins which interact with mutant VEGF proteins is the most elegant and specific means to correct these molecular defects. In addition, some genetic changes may result in altered conformational states of the receptor. This, in turn, may functionally inactivate the tyrosine kinase activity, making it incapable of signal transduction. In order to restore the function of such mutant proteins an antibody may be used which binds to an epitope and induces a conformational change of the protein thereby restoring the wild type function. Thus, the methods of the invention are also useful to screen e.g.,



antibody, Fab, Fv or scFv expression libraries wherein the DNA sequence encoding said antibodies or derivatives thereof are under the control of the regulatory sequence of the invention. It is, of course, evident to the person skilled in the art that also other protein or peptide expression libraries using the regulatory sequences of the invention may be employed.

Further, the present invention relates to the use of the recombinant DNA molecule, vector, cell, pharmaceutical compositions, diagnostic compositions or a transgenic non-human animal of the invention for the identification of a chemical and/or biological substance capable of suppressing or activating and/or enhancing the transcription, expression and/or activity of genes and/or its expression products in endothelial cells.

In a preferred embodiment, the chemical or biological substance used in the methods and uses of the present invention is selected from the group consisting of peptides, proteins, nucleic acids, antibodies, small organic compounds, hormones, neural transmitters, peptidomimics, and PNAs (Milner, Nature Medicine 1 (1995), 879-880; Hupp, Cell 83 (1995), 237-245; Gibbs, Cell 79 (1994), 193-198).

The present invention further relates to a method of inhibiting a vascular disease in a subject, comprising contacting an artery of said subject with the recombinant DNA molecule or vector of the invention, wherein said protein reduces or prevents the development of the vascular disease, preferably said protein reduces proliferation of smooth muscle cells.

In a further embodiment the present invention relates to the use of a recombinant DNA molecule, vector, nucleic acid molecule of the invention and/or substance identified by a method of the invention for the preparation of a composition for directing and/or preventing expression of genes specifically in endothelial cells and/or for the preparation of a pharmaceutical composition for treating, preventing and/or delaying a vascular disease and/or a tumorous disease in a subject. The

upregulation and activation of the Flk-1 receptor in peri-tumoral endothelial cells is believed to be involved in the neovascularization of various human or experimental tumors (Plate, 1994; Ferrara, Curr. Opin. Nephrol. Hypertens. 5 (1996), 35-44). This hypothesis is supported by experiments in which the inhibition of Flk-1-mediated signal transduction strongly inhibits tumor angiogenesis and tumor growth (Millauer, Nature 367 (1994), 576-579; Millauer, Cancer Res. 56 (1996), 1615-1620). Thus, by using compounds of the present invention described above capable of inhibiting Flk-1 gene expression, it is possible to ameliorate tumorous diseases which depend on the expression of the FLK-1 gene.

In a further embodiment, the present invention relates to the use of a recombinant DNA molecule, vector and/or the nucleic acid molecule of the invention for the preparation of a pharmaceutical composition for inducing a vascular disease in a non-human animal or in a transgenic non-human animal described above.

In a preferred embodiment of the methods and uses of the invention, the vascular disease is atherosclerosis and/or a neuronal disorder. Further possible methods and uses in accordance with the present invention will be evident to the person skilled in the art and are described in, for example, WO 95/13387, WO 94/11499 and WO 97/00957.

The recombinant DNA molecules, vectors, nucleic acid molecules, compounds, uses and methods of the invention can be used for the treatment of all kinds of disorders and diseases hitherto unknown as being related to or dependent on the modulation of genes specifically expressed in the endothelium. The recombinant DNA molecules, vectors, nucleic acid molecules, compounds, methods and uses of the present invention may be desirably employed in humans, although animal treatment is also encompassed by the methods and uses described herein. Thus, the present invention provides for the use of a regulatory sequence as defined above for enhancing and/or directing gene expression in endothelial cells in any kind of organism.

These and other embodiments are disclosed and encompassed by the description and examples of the present invention. Further literature concerning any one of the methods, uses and compounds to be employed in accordance with the present invention may be retrieved from public libraries, using for example electronic devices. For example the public database "Medline" may be utilized which is available on the Internet, for example under <http://www.ncbi.nlm.nih.gov/PubMed/medline.html>. Further databases and addresses, such as <http://www.ncbi.nlm.nih.gov/>, <http://www.infobiogen.fr/>, [http://www.fmi.ch/biology/research\\_tools.html](http://www.fmi.ch/biology/research_tools.html), <http://www.tigr.org/>, are known to the person skilled in the art and can also be obtained using, e.g., <http://www.lycos.com>. An overview of patent information in biotechnology and a survey of relevant sources of patent information useful for retrospective searching and for current awareness is given in Berks, TIBTECH 12 (1994), 352-364.

The figures show:

**Figure 1:** Nucleotide sequence of the murine Flk-1 gene. The ATG codon is at position +299. The three exons are indicated in bold. Motifs for transcription factors are underlined. VRE: vascular response element,

**Figure 2:** Map of reporter gene construct pGL2-B. Arrows symbolize functional elements. Luc: luciferase gene, AMP: ampicillin resistance gene, f1ori: replication origin for bacteriophage f1.

**Figure 3:** Map of reporter gene construct pGLacZ. Arrows symbolize functional elements. LacZ:  $\beta$ -galactosidase gene, AMP: ampicillin resistance gene, f1ori: replication origin or bacteriophage f1.

**Figure 4:** Partial structure and functional analysis of the mouse Flk-1 locus. A) Restriction enzyme map of the region encompassing the first three

exons (represented by hatched boxes). Subfragments containing parts of intron 1 or intron 2 are indicated. Abbreviations for restriction enzymes are: B, BamHI, Xh, XhoI, SI, Sall. B and C, luciferase reporter gene assays of various constructs following transient transfection of bovine aortic endothelial cells. B) Transfection assay of the intron fragments in combination with the Flk-1 promoter region of bp -640 to bp +299. The values were coordinated with 5'-In1 fragments with respect to the activity of the construct. C) The intron fragments were tested in combination with a 4.4 kbp Flk-1 promoter fragment spanning the region from -4.1 kbp to +299 bp of the Flk-1 gene. NIH 3T3 cells were used as a reference for non-endothelial cells. RLU, relative light units

**Figure 5:** Analysis of the intron in transgenic mice. The embryo (10.5 days) was stained overnight with X-Gal. The reporter gene was under the control of the intron enhancer (3'-In1, cf. Fig. 4A) and of the Flk-1 promoter fragment ranging from nucleotides -4100 to +299. A) Top lateral view. B) Top dorso-cranial view.

**Figure 6:** Histological evaluation of a transgenic embryo. The embryo shown in Fig. 5 was embedded in paraffin. The cuts were stained with neutral red. A) Pseudo transversal cut through the head region. B) Magnification of a section from A. C) Pseudo transversal cut of a caudal region. 1: 4th ventricle of cerebrum, 2: acoustic vesicle/otocyte, 3: V. cardinalis anterior, 4: third ventricle of cerebrum, 5: endbrain vesicle, 6: optic vesicle, 7: ganglion trigeminale (V), 8: chorda dorsalis.

**Figure 7:** Functional analysis of the first two introns of the Flk-1 gene in vivo. The depicted embryo (10.5 days) carries the  $\beta$ -galactosidase gene under the control of the Flk-1 promoter (-4.1 kbp to +299 bp) and the

first 6.2 kbp of the transcribed region (cf. Fig. 4A). The staining was carried out as described in Fig. 5.

- Figure 8:** In vivo characterization of the intron enhancer in combination with the strongest promoter fragment. All three embryos carry the  $\beta$ -galactosidase gene under the control of the Flk-1 promoter fragment of bp -640 to bp +299 and the intron enhancer. The staining was carried out as described in Fig. 5.
- Figure 9:** Detailed analysis of the left-hand embryo from Fig. 8A) Left lateral view. B) Sectional magnification of A. C) Right lateral view.
- Figure 10:** Histological evaluation of the embryo depicted in Fig. 9. The embryo was embedded in paraffin and was cut into 10  $\mu$ m slices. The cuts were stained with neutral red. A) Pseudo transversal cut through the head region. B) Magnification from a similar cut level as in A. C) Pseudo transversal cut from a more caudally located section. D) Pseudo transversal cut from thoracic section. 1: 4th ventricle cerebrum, 2: 3rd ventricle cerebrum, 3: endbrain vesicle, 4: A. carotis interna, 5: ganglion trigeminale (V), 6: V. cardinalis anterior, 7: neural tube, 8: esophagus, 9: V. cardinalis posterior, 10: aorta dorsalis, 11: endocardium of the heart atrium, 12: vessels of the myocardium.
- Figure 11:** Reporter gene analysis of Flk-1 gene regulatory elements in transgenic mouse embryos. The lacZ reporter gene was fused to regulatory elements derived from the mouse Flk-1 gene and tested for  $\beta$ -galactosidase expression in transgenic mouse embryos. A) 10.5 day transgenic mouse embryo expressing lacZ under the control of a 939 bp promoter fragment in combination with a 2.3 kbp XhoI/BamHI fragment of the first intron spanning the region from +1677 bp to +3947 bp of the *Flk-1* gene. This embryo was derived from a foster

mother. Most if not all developing vascular structures show  $\beta$ -galactosidase expression, for example the endocardium of the heart, the dorsal aorta, intersomitic vessels or vessels of the developing brain. B) 11.5 day embryo of a transgenic mouse line that was established with the same construct. C) An 11.5 day *Flk-1/lacZ* knock-in embryo in which the *lacZ* gene is expressed from the endogenous *Flk-1* locus shows a highly similar staining. However, note the absence of  $\beta$ -galactosidase expression in small blood vessels of the yolk sac. D-F) Paraffin sections of the  $\beta$ -galactosidase stained embryo from (B) demonstrate  $\beta$ -galactosidase expression in the paired dorsal aortae (D), a venous vessel connected with the heart (E), and capillaries invading the neural tube (F). G)  $\beta$ -galactosidase expression in a transgenic embryo containing the *tk* promoter in combination with the 2.3 kbp *XhoI/BamHI* fragment of the *Flk-1* first intron. H)  $\beta$ -galactosidase expression in a transgenic embryo containing a construct with *Flk-1* promoter sequences (-640 bp/+299 bp) in combination with the 510 bp *SwaI/BamHI* fragment of the first intron spanning the region from +3437 bp to +3947 bp of the *Flk-1* gene. Bar D) –F) 100  $\mu$ M.

**Figure 12:** Nucleotide sequence of the *Flk-1* intron enhancer and putative transcription factor binding sites. Sequences matching known transcription factor binding sites are underlined. This sequence is deposit in the GeneBank database (accession number AF061804).

**Figure 13:** Analysis of transgene expression during early development and in newborn mice. The transgenic mouse line 2603 expresses *lacZ* under the control of the *Flk-1* promoter (-640 bp/+299 bp) in combination with the 2.3 kbp *Flk-1* intron enhancer A) Frontal view on a whole mount  $\beta$ -galactosidase stained 7.8-day embryo. The arrow indicates transgene expression in the extraembryonic mesoderm. B) and C)

Paraffin sections from the embryo shown in A demonstrate transgene expression in endothelial cells of the allantois (B) and the yolk sac (C). D-H, LacZ staining of spleen (D), kidney (E), lung (F), liver (G) and thymus (H) from a postnatal day 5 transgenic mouse. EM, extraembryonic mesoderm. Bars, 25  $\mu$ M (C), 100  $\mu$ M (B,D,E,F,G,H).

**Figure 14:** The 5'-UTR is required for expression of the *Flk-1* gene in the yolk sac vasculature. Transgenic mouse embryos that contain a *Flk-1* promoter and 5'UTR (-640 bp/+299 bp) / enhancer (+1677 bp/+3947 bp) reporter gene construct show a complete vascular expression in the yolk sac vasculature (A and B). In contrast, the yolk sac of *Flk-1/lacZ* knock-in embryos which lack part of the 5'UTR show expression only in large collecting vessels that connect with the embryo, but not in the smaller vessels (C). Bar, 500  $\mu$ M.

**Figure 15:** HIF-2 $\alpha$  stimulates *Flk-1* gene expression. A293 cells were co-transfected with a reporter gene construct containing *Flk-1* promoter sequences from bp -640 to bp +299 and with expression vectors encoding the murine HIF-1 $\alpha$  and HIF-2 $\alpha$  cDNAs, respectively. Relative promoter activities were determined as described in Materials and Methods. The promoter activity of the control transfection was arbitrarily set to 1.

The examples illustrate the invention.

#### **Example 1: Cloning and construction of *Flk-1* intron/reporter gene vectors**

DNA clones containing the 5' region of the mouse *Flk-1* gene were isolated from a library prepared from 129/SvJ mice in  $\lambda$  Dash II vector (Stratagene) (Röncke, supra) or in  $\lambda$  FIX II or obtained from the P1 Library (Genome Systems, St. Louis). A 21 kb region of the mouse *Flk-1* gene, contained in the DNA insertions of two  $\lambda$  phages 6

and 16, including approximately 15 kb of 5' flanking sequences, exons 1, 2 and 3, and introns 1 and 2 was characterized by restriction enzyme mapping and Southern blot analysis. Lower DNA fragments of the phage clones were cloned into pBluescript vector DNA (Stratagene) and used for further characterization. Sequencing was performed using an automatic Sequencer (373A, Applied Biosystems). The nucleotide sequence of the *Flk-1* intron enhancer is deposited in the Genbank database (accession number AF061804). The search for potential transcription factor binding sites was performed with the MatInspector software (Quandt, (1995) Nucl. Acids Res. 23, 4878-4884).

The DNA sequence (SEQ ID NO: 1) of a 12.8 kb region spanning from about position -6,660 kb (relative to the transcriptional start site) to approximately position +6,135 kb (located in the third exon) was determined (Fig. 1). Figure 4A shows a schematic representation of the first 6.5 kbp of the transcribed region of the murine *Flk-1* gene. Exons I, II and III are emphasized as hatched boxes. The first intron having a length of 3.6 kbp is subdivided into two regions (5'-In1 and 3'-In1). The region In-2 contains the entire second intron, the second exon, the 3' end of the first intron and part of the third exon. This subdivision into various intron fragments was maintained in the following analyses. The reporter gene constructs used were derived from pGL2 basic vector (Promega) that contains a promoterless luciferase gene. Luciferase reporter gene constructs were generated for transfection of cells in vitro. For use in transgenic mice in vivo, plasmids were used in which the luciferase reporter gene was replaced by a lacZ reporter gene.

In order to generate (luciferase) reporter gene constructs, *Flk-1* promoter fragments were amplified by PCR and cloned into pGL2 (Promega) vector DNA 5' to the luciferase gene as described by Röncke, supra; see also Figure 2. In short, the upstream primers used were -1900: 5'-GGG GTA CCG AAT TCT AAA TGG GGC GAT TAC C-3' (SEQ ID NO 2); -640: 5'-GTG GTA CCC AAA CAC TCA ACA CCA CTG-3' (SEQ ID NO: 3); -624, 5'-TCG GTA CCG ACC CAG CCA GGA AGT TC-3' (SEQ ID NO: 4); the downstream primer was +299, 5'-TTG CTA AGC TTC CTG CAC CTC GCG CTG GG-3' (SEQ ID NO: 5). To generate the construct ranging from -4100 to +299, a HindIII-EcoRI fragment of recombinant lambda phage 6 from P1



Library (Genome Systems, St. Louis) was inserted into the plasmid ranging from -1900 to +299. Vectors that contained Flk-1 intron sequences in addition to promoter sequences were generated as follows: specific intron sequences were amplified by PCR from cloned Flk-1 genomic DNA and inserted downstream of the reporter gene. Primers used for amplification were 5'-In1down: 5'-AGG GAT CCA CTC TTT AGT AGT AAG GCG-3' (nucleotides 7036-7057 of SEQ ID NO: 1, SEQ ID NO: 6); 5'-In1up: 5'-ACC TCG AGA CTT GGA TGG CAC-3' (nucleotides 8324-8342 of SEQ ID NO: 1, SEQ ID NO: 7); 3'-In1down: 5'-GGG CTA TAA TTG GTG CCA TCC-3' (nucleotides 8312-8332 of SEQ ID NO: 1, SEQ ID NO: 8); 3'-In1up: 5'-GGA TGG AGA AAA TCG CCA GGC-3' (nucleotides 10637-10658 of SEQ ID NO: 1, SEQ ID NO: 9); IN2A: 5'-GTG TGC ATT GTT TAT GGA AGG G-3' (nucleotides 10571-10593 of SEQ ID NO: 1, SEQ ID NO: 10); IN2B: 5'-CAT AGA CAT AAA CAG TGG AGG C-3' (nucleotides 12849-12871 which is part of the cDNA sequence published by Millauer (1993), supra, SEQ ID NO: 11). For the subsequent experiments the vector indicated in Fig. 3 was used. It represents a modification of the pGL2 basic vector in which the corresponding Flk-1 promoter fragments were inserted into the KpnI and HindIII restriction sites of the polylinker (Fig. 2). Also the luciferase reporter gene was replaced by the  $\beta$ -galactosidase gene (Schlaeger, Proc. Natl. Acad. Sci. USA 94 (1997), 3058-3063). For an analysis of the intron fragments were cloned into the BamHI and Sall restriction sites indicated. DNA manipulations, PCR amplification and DNA sequencing were performed according to conventional methods known in the art as described, for example in Sambrook, supra and PCR Technology, Griffin and Griffin, eds., RC Press London (1994).

### Example 2: Functional analysis of the intron of the Flk-1 gene in vitro

Figure 4 shows the result of transient transfections in BAECs. The corresponding intron fragments were combined with a Flk-1 promoter fragments which comprised nucleotides -640 to +299. The promoter activity was standardized with respect to the promoter activity of the construct containing the 5'-In1 fragment.

Tissue culture and transient transfections were performed as follows:

All cells were cultured in DMEM+ supplemented with 10% FCS (Sigma) and as described in Schaefer, 1997. bEnd5 cells were generated by transformation with the Polyoma middle-T oncogene as described earlier (Montesano, Cell 62 (1990), 435-445). Bovine aortic endothelial cells (BAECs) were prepared as described (Schwartz, In Vitro 14 (1978), 966-980). NIH 3T3, C2C12 and L cells were obtained from ATCC. Transient transfections were performed using the CaPO<sub>4</sub>-precipitation method according to Chen and Okayama (Mol. Cell Biol. 7 (1987), 2745-2752), optionally with modifications as described (Röncke, 1996). The transfection efficiency was monitored by co-transfection of a  $\beta$ -galactosidase reporter vector. Each construct was transfected at least six times in three independent experiments. Cells were grown to 70% confluence in 6-cm dishes prior to transfection. Cells were washed 16 hrs after addition of CaPO<sub>4</sub>-precipitate and incubated for further 48 hrs. In each experiment, 6  $\mu$ g luciferase and 1  $\mu$ g pCMV5 (Röncke, supra) lacZ reporter gene constructs were used. Cells were lysed in 1 x reporter-lysis-buffer (Promega) for 15 min on a test tube-rotator. After centrifugation, the supernatant was transferred to a fresh tube and stored at -80 °C or taken for luciferase and lacZ assay immediately. Reporter-gene assays for  $\beta$ -galactosidase activity were performed according to Eustice (Biotechniques 11 (1991), 739-740). Chlorophenol red- $\beta$ -D-galactopyranoside (CPRG) was used as a substrate and the conversion was measured at 575 nm in an ELISA-reader (Biometra). Extracts were diluted to obtain OD<sub>575nm</sub> values between 0.2 and 0.8. These values were used to standardize for transfection efficiency after subtracting a background value, as determined from a cell extract of a transfection without lacZ-reporter plasmid but with a luciferase-reporter plasmid. Luciferase-reporter gene assays were performed with the same extracts as described by the manufacturer (Promega). Luciferase activity was measured with a luminometer (LB96P, Berthold) and calculated as per cent of the activity of the pGL2-promoter plasmid (Promega).

Construct	5'-In1	3'-In1	In2
BAEC	100+/-0%	128+/-34%	136+/-52%
3T3	100+/-0%	55+/-15%	74+/-33%

**Table I:** Functional analysis of the intron of the Flk-1 gene. The upper line indicates the corresponding intron fragment which was analyzed in combination with the Flk-1 promoter (-640 bp/+299 bp).

Figure 4C shows the results of another transfection assay of the intron fragments. It was carried out as described above, with the exception that a Flk-1 promoter fragments was used that comprised the region between nucleotides -4100 and +299. Also, a fragment was analyzed that contained the entire first intron, the second exon, the second intron and part of the third exon shown in Figure 4A.

Construct	5'-In1	3'-In1	In2	In1+2
BAEC	100+/-0%	206+/-81%	119+/-51%	154+/-68%
3T3	100+/-0%	71+/-32%	85+/-27%	35+/-12%

**Table II.:** Functional analysis of introns of the Flk-1 gene. The upper line indicates the corresponding intron fragment which was analyzed in combination with the Flk-1 promoter (-4100 Bp/+299 Bp).

An analysis of this experiment revealed that the construct with the 3' region of the first intron in BAECs had an activity that was twice that of that containing the 5' region of the first intron. Also, it showed 85% higher activity than the construct with the second intron ( $p=0.0153$ ). The 4.5 kbp longer construct In1+2 that also

contained the 3' region of the first intron, too, revealed an activity that was markedly higher in BAECs than in 3T3 cells.

A functional analysis of the first 6.5 kbp of the transcribed region of the murine Flk-1 genes lead to the identification of an endothelial-specific positive regulatory element. This regulatory sequence is located in the region between the XhoI and BamHI restriction site in the first intron of the Flk-1 gene (cf. Fig. 4A). It is functional in both orientations since the intron is used in an antiparallel manner with respect to the Flk-1 promoter fragment in the construct referred to as 3'-In 1. In construct In1+2, however, the original orientation was maintained. A sequence analysis of the intron enhancer lead to the identification of two potential GATA binding sites at position +1927 bp and +3514 bp; (Evans, Proc. Natl. Acad. Sci. USA 85 (1988), 5976-5980; Orkin, Blood 80 (1992), 575-581), a potential AP-1 binding site at position +2210 bp; (Lee, Cell 49 (1987), 741-752) and two PEA3 consensus sequences at position +3494 bp and +3741 bp; (Martin, Proc. Natl. Acad. Sci. USA 85 (1988), 5839-5843).

### **Example 3: Functional characterization of the Flk-1 promoter in vivo.**

So far, analyses of the murine Flk-1 promoter have been restricted to in vitro systems (Rönicke, supra; Patterson, supra). The investigation of the promoter activity in vitro is an important tool in promoter characterizing since it is useful to assay a large number of promoter constructs for their activity in a short time. However, this situation is always an artificial one since not all factors that are relevant in vivo can also be reconstituted in vitro. While an in vitro investigation of a promoter yields important information on the mechanisms of gene regulation it is only the in vivo characterization that can yield the final proof for the relevance of the elements identified. An excellent test system for promoter analysis in vivo are transgenic mice. In this model the corresponding promoter fragment was cloned before a reporter gene, isolated together with this reporter gene and injected into fertilized mouse oocytes. In many cases, successful integration of the promoter reporter construct into the mouse genome lead to transgenic mice which contain the

construct in every cell. This test system, in addition to the analysis of the promoter activity during embryonic development and in the adult animal, allows a tissue-specific characterization of the promoter activity.

For the investigation of the Flk-1 promoter in transgenic mice the bacterial  $\beta$ -galactosidase reporter gene was chosen since the gene product is easily detectable by color reaction and remains at the location of production due to its limited solubility. In this manner it is possible to identify cells in which the corresponding Flk-1 promoter fragment since only there an expression of  $\beta$ -galactosidase took place.

When producing transgenic mice it was taken care that no regions originating from the vector were injected along with the promoter. First, Flk-1 promoter fragments comprising the regions between nucleotides -640 and +299, -1900 and +299 as well as -4100 and +299 were investigated. The constructs were based on plasmid pGL-2B described in Figure 2 with the exception that the luciferase reporter gene was replaced by the  $\beta$ -galactosidase gene. All injection fragments used in the examples were obtained by restriction digestion with the enzymes KpnI and SalI. Transgenic mice were generated as described by (Hogan, Manipulating the Mouse Embryo (1994), Cold Spring Harbor Laboratory Press, New York). Fertilized oocytes were isolated from superovulated C57BL/6 x C3H/He F1 mice, microinjected and reimplanted into pseudopregnant females of the same hybrid-mouse strain. Mice were sacrificed at day 10.5 or 11.5 of gestation, and embryos were analyzed by whole mount LacZ staining for transgene expression. The embryos to be examined were isolated on day 10 after reimplantation of the injected oocytes. Analysis of the transgenic embryos revealed that although promoter activity could be detected, none of the constructs was capable of conferring reproducible expression of the reporter gene in the endothelium.

#### **Example 4: Functional characterization of the Flk-1 intron in vivo**

After analysis of the Flk-1 promoter region from -4.1 kbp to +299 Bp the intron which was identified in vitro was then examined for its function in vivo. For this purpose, a construct similar to that shown in Figure 3 was used which contained an Flk-1

promoter fragment ranging from nucleotide -4100 to base pair +299 and the intron enhancer (3'-In1, cf. Fig. 4A). The staining and fixation of the embryos was performed as follows: The mid-day of the plug observation was counted as E0.5. The embryos were dissected out in ice-cold PBS and fixed in ice-cold 2 % (w/v) paraformaldehyd, 2 mM  $MgCl_2$ , 2 mM EGTA, 0.1 M Pipes buffer, pH 6.9 for 15 to 120 minutes. The embryos were rinsed with PBS three times for 5 minutes each. The LacZ expression was detected by incubating the embryos at 30 °C overnight in 0.1 % X-gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside), 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide 5 mM, 1 to 2 mM magnesium chloride, 0.002 to 0.02 % NP-40, 0.01 % or 0.25 mM sodium deoxycholate, PBS, pH 7.0. After the staining, embryos were rinsed in PBS and postfixed at 4 °C overnight in 2 % paraformaldehyde, 0.1 % glutaraldehyde, PBS, pH 7.0. For whole-mount photography, the postfixed embryos were rinsed in PBS and equilibrated for 3 minutes each in (optionally 30 %) 50 % glycerol and then in 70 % glycerol. Figure 5 shows an embryo (embryonic day 10.5) which was isolated after injection of the fragment. In Figure 5A a color reaction in vessels of the developing brain can be clearly discerned. Also superficial vessels in the body's middle (dorsal) and a staining in the liver bud can be observed. Figure 5B shows the dorsal, caudal region of the same embryo. It proves that the vessels in both halves of the head were stained.

For an exact localization of the stained cells the embryo was embedded in paraffin, cut into slices of 10  $\mu$ m and counterstained with neutral red. Cryostat sectioning and lacZ staining of organs from postnatal mice was performed as described (Schlaeger, 1997).

The results of the histological analysis are shown in Figure 6. It shows pseudo transversal cuts of the embryo. In section 6A a staining of the inner lining of the V. cardinalis anterior (3) and of other superficial vessels can be seen. Figure 6B represents a strong magnification of a section of 6A with the staining of endothelial cells within the V. cardinalis anterior. Figure 6C shows a more caudally located region. Again, the staining of the V. cardinalis anterior and of superficial vessels with

wide lumen and thin walls as well as of vascular structures in the neural tube is clearly visible. Also, a staining of the chorda dorsalis (9) can be observed. However, in none of the cases a staining of arterial vessels could be observed.

The subsequent injection of the same fragment lead to a total of eight further transgenic embryos which displayed an identical expression pattern albeit in two cases of weaker nature. Thus, the intron enhancer exhibited in vivo an effect that was even more marked than in vitro. In combination with a promoter fragment which on its own had a very variable expression pattern it ensures a reproducible expression pattern with clear endothelium specificity, however, covering a substantial part of the endogenous Flk-1 expression pattern.

#### **Example 5: Functional analysis of the introns I and II of the murine Flk-1 gene in vivo**

Since the intron enhancer in combination with an Flk-1 promoter fragment displayed an endothelium-specific function in transgenic mice covering a substantial part of the endogenous expression pattern, the further search for in vivo relevant, gene regulatory elements was extended to other intron regions. For this purpose, the construct containing the promoter region between nucleotides -4100 to +299 and the first 6.5 kbp of the transcribed region (In1+2; cf. Fig. 4) was used. Figure 7 shows an embryo on embryonic day 10.5 which was obtained after injection of this fragment. Again, a staining of the vessels in the developing brain as well as superficial vessels the of the liver bud was visible. The following injections yielded four further transgenic embryos which displayed the same pattern. A combination of the promoter region used with only the 5' end of the first intron (5'-In1; cf. Fig. 4), however, yielded no endothelium-specific expression pattern.

**Example 6: Combination of the intron enhancer with the Flk-1 promoter fragment that was the most potent in vitro**

To investigate whether the repressing elements of the murine Flk-1 promoter between nucleotides -4100 and -640 are functional also in combination with the intron enhancer, a shorter construct without these inhibitory regions was used for further analysis. It contains the intron enhancer (3'-In1) and the 5' region from base pair -640 to nucleotide +299. This 5' region displayed the highest activity in vitro. Figure 8 shows three transgenic embryos (embryonic day 10.5) which were obtained after injection of the fragment. All three display a more marked staining in vascular structures than the embryos analyzed so far. While the embryo on the right hand shows a weak staining, the left-hand embryo yields a very strong expression in virtually all vessels. The embryo in the middle holds a medium position as regards the completeness of its expression pattern, i.e., it lacks expression in the heart although it resembles strongly the embryo on the left hand as regards the staining of the other structures. In Figure 9A the left-hand embryo from Fig. 8 is shown in more detail. The strong staining of the heart in the region of the atrium and ventricle is particularly clearly visible. Furthermore, the vessels of the developing brain, the vessels between the somites, the aorta dorsalis as well as the fine capillary plexus on the body's surface are stained. Figure 9B shows a sectional magnification of 9A. Here, the staining of the vessels in the head region as well as the expression in the superficial capillary plexus is visible. In Figure 9C the same embryo is shown from the other side. In addition to the structures described in Figure 9A also a staining of the chorda dorsalis can be observed.

The embryo shown in Figure 9 was embedded in paraffin and cut to slices. The cuts dyed with neutral red are shown in Fig. 10. Figure 10A shows a pseudo transversal cut through the head region. Particularly prominent is the branching of the A. carotis interna (4) in addition to the staining of other vascular structures. Figure 10B represents a magnification of a similar cut; here, too, the branching of the A. carotis interna is particularly striking. Figure 10C shows a more caudal cut which in terms of its position roughly corresponds to the cut shown in Figure 6A. Here, however, in



addition to the staining of the V. cardinalis anterior (6) an expression in the branching of the A. carotis interna (4) and other vascular structures is visible. Figure 10D represents an even more caudally located region. A staining in the venous endothelium (V. cardinalis posterior, 9) and in the arterial structures (aorta dorsalis, 10) can be observed. Furthermore, the endocardium of the atrium as well as the vessels in the trabeculae of the heart ventricles display an expression.

A total of seven transgenic embryos was analyzed after injection of this fragment (-640 bp/+299 bp/3'-In1). Safe for one which showed no staining, all embryos displayed an expression of the  $\beta$ -galactosidase in endothelial structures. The staining was regularly more marked than in combination with the negative regulatory elements between nucleotides -4100 and -640. Thus, the in vitro identified regions displayed a function in vivo. The deletion of these negative regulatory elements yielded a construct that lead to a reproducible expression in venous and arterial endothelium.

#### **Example 7: Endothelium-specific expression mediated by Flk-1 regulatory sequences in vivo**

When the Flk-1 promoter fragment with the strongest in vitro activity (-640 bp/+299 bp; Röncke, supra) was tested in combination with the 2.3 kbp XhoI/BamHI fragment of the first intron that showed endothelium-specific activity in vitro (3'-In1; +1677 bp/+3947 bp; see Example 6), a reproducible vascular lacZ expression in transgenic E10.5 mouse embryos derived from foster mothers was observed (Table III). In these embryos, the lacZ reporter gene was expressed in developing vascular structures, such as capillaries in the head region, intersomitic vessels, the dorsal aorta, and in the heart anlage (Fig. 11A). Sectioning of these embryos confirmed that the  $\beta$ -galactosidase protein was confined to vascular endothelium. This in vivo analysis demonstrated that the intron sequences in combination with the Flk-1 promoter confer an endothelium-specific expression pattern that closely resembles the expression pattern of the endogenous Flk-1 gene (Millauer, 1993). Moreover, the intron fragment could also direct endothelial cell-specific lacZ expression when used

in an inverted orientation in the reporter construct (Construct -640 bp/+299 bp//+3947 bp/+1677 bp; see Table III).

**Table III.** Summary of the in vivo activity of different *flk-1* constructs

Construct	TG	ES	ET	NO
-4100/+299	11	0	3	8
-1900/+299	31	1	10	20
-640/+299	3	0	1	2
-640/+299 // 3'Intron +1677/+3947	7	6	0	1
-640/+299 // 3'Intron +3947/+1677	4	3	1	0
-640/+299// 3'Intron +3437/+3947	7	5	0	2
-640/+254 // 3'Intron +1677/+3947	12	8	1	3
tk // 3'Intron +1677/+3947	15	3	0	12
-5500/+299 // Intron I+II	3	2	0	1

Embryos transgenic for the constructs given above were generated, and LacZ staining and genotyping was performed at E10.5 or E11.5 as described in Example 4. Constructs are defined by the position of the promoter or intron fragments in bp relative to the transcription initiation site of the endogenous *Flk-1* gene. TG, number of transgenic embryos; ES, number of embryos showing endothelial-specific staining; ET, number of embryos showing ectopic staining; NO, number of embryos showing no staining at all.

Transgenic mouse lines were generated with this reporter gene construct (-640 bp/+299 bp//+1677bp/+3947bp) containing the *Flk-1* regulatory sequences. One of these lines (2603) showed a complete vascular expression of the reporter gene at

E11.5 and was analyzed further (Fig. 11B). Sectioning of  $\beta$ -galactosidase stained E11.5 transgenic embryos revealed that reporter gene expression was confined to the endothelium of blood vessels, e.g. in the endothelium of the dorsal aorta (Fig. 11D), in venous vessels (Fig. 11E) and in the perineural vascular plexus and sprouting capillaries invading the neural tube (Fig. 11F). To determine whether transgene expression in this mouse line reproduced the complete expression pattern of the endogenous *Flk-1* gene, the lacZ staining pattern of these embryos was compared to heterozygous *Flk-1* mutant mouse embryos which express the lacZ gene from the endogenous *Flk-1* locus (Shalaby, Nature 376 (1995), 62-66). In these knock-in mice, the lacZ gene was inserted into the endogenous *Flk-1* locus via homologous recombination and is therefore expected to reproduce the expression pattern of the endogenous *Flk-1* gene (Shalaby, 1995). The lacZ staining pattern of transgenic embryos and the knock-in embryos at E11.5 was indistinguishable (Fig. 11B,C). It is concluded from these data that the -640 bp/+299 bp promoter region of the *Flk-1* gene and the 2.3kbp XhoI/BamHI fragment of the first intron contain regulatory elements that are sufficient for endothelial-specific gene expression in developing mouse embryos.

**Example 8: The first intron of the *Flk-1* gene contains an autonomous endothelium-specific enhancer**

To assess the role of the 2.3 kbp XhoI/BamHI fragment of the first *Flk-1* intron in endothelium-specific gene expression, it was further investigated whether the intron sequences can confer endothelium-specific expression to the heterologous herpes simplex virus-thymidine kinase (*tk*) promoter. This promoter has no intrinsic endothelial cell specificity (Schlaeger, 1997). A lacZ reporter gene construct was generated that contained the *tk* promoter, in combination with the 2.3 kbp BamHI/XhoI fragment of the first intron (+1677 bp/+3947 bp). The *tk* promoter sequences were amplified from the plasmid ptkSDKlacZ (Schlaeger, 1997) using oligonucleotides tk5' (5'-CCGGTACCCAAACCCGCCAGCGTCTTG-3'; SEQ ID

NO: 16) and tk3' (5'-CCGACAAGCTTGGTCGCTCGGTGTTTCGAGG-3'; SEQ ID NO: 17). The PCR product was digested with KpnI and HindIII.

From the  $\beta$ -galactosidase reporter construct described in Example 2 (Fig. 2), *Flk-1* promoter sequences were excised and removed by KpnI and HindIII digestion. The *tk* promoter was then subcloned in the KpnI and HindIII restriction sites of the vector. Transgenic mouse embryos generated with this construct showed vascular reporter gene expression (Fig. 11G). The  $\beta$ -galactosidase staining observed in these embryos was weaker than in embryos expressing *lacZ* under the control of the -640 bp/+299 bp *Flk-1* promoter in combination with the intron fragment (Fig. 11A,B). Also, the frequency of transgenic mouse embryos expressing this transgene was significantly reduced, when compared with constructs driven by the -640 bp/+299 bp *Flk-1* promoter in combination with the intron fragment (Table III). This indicates that the *tk* promoter lacks positive acting elements which are present within the *Flk-1* promoter. However, these results show that the *Flk-1* intron fragment alone, in contrast to the *Flk-1* promoter, can reproducibly target reporter gene expression to the endothelium. Taken together, the results of both the in vitro and in vivo experiments in this study demonstrate that sequences located in the first intron of the mouse *Flk-1* gene act as an autonomous endothelium-specific enhancer.

In order to further characterize the minimal intron sequences that are required for endothelium-specific expression, we analyzed whether shorter intron fragments were also active in combination with the 939 bp promoter region of the *Flk-1* gene (-640 bp/+299 bp). By this deletion analysis, the intron enhancer was localized to a 510 bp *SalI/BamHI* fragment (+3437 bp/+3947 bp) which is located immediately upstream of the second exon. This fragment was sufficient to stimulate endothelium-specific *lacZ* expression in transgenic mouse embryos (Fig. 11H, Table III). The DNA sequence of this fragment (Fig. 12) contains potential binding sites for the GATA and Ets transcription factors, and for *Scf/Tal-1*, all of which have been implicated to play a role in angiogenesis (reviewed by Risau, Nature 386 (1997), 671-674). Whether these consensus sequences represent functional transcription factor binding sites remains to be determined.

### **Example 9: *Flk-1* regulatory sequences target endothelium-specific transgene expression throughout development**

To test whether the regulatory sequences of the *Flk-1* promoter and enhancer identified can reproduce the endogenous *Flk-1* expression pattern throughout development, the *lacZ* expression pattern of the transgenic mouse line 2603 (Fig. 11B) was further analyzed at various stages of embryonic development and at postnatal days 5 (P5) and 120 (P120). In this mouse line, the transcription of *lacZ* is driven in combination by the -640 bp/+299 bp *Flk-1* promoter and the 2.3 kbp BamHI/XhoI intron enhancer fragment. The earliest stage during which transgene expression was detectable by whole mount LacZ staining was in E7.8 embryos (Fig. 13A). This is the earliest stage that was examined. The analysis of sections of these embryos confirmed that the transgene was expressed in angioblasts of the allantois and the yolk sac (Fig. 13B,C). Moreover, transgene expression was restricted to the vascular endothelium at all stages of embryonic development examined. To determine if the transgene expression persists after birth, we performed lacZ staining of cryostat sections from several different organs of P5 and P120 transgenic mice. Strong LacZ staining was detected in vessels of the spleen, kidney, thymus, liver and lung from P5 animals (Fig. 13D-H). However, *lacZ* expression was downregulated in most vascular beds of P120 animals, as is the case for the endogenous *Flk-1* (Millauer, 1993; Kremer, Cancer Res. 57 (1997), 3852-3859). Taken together, these results support the conclusion that the identified *Flk-1* regulatory sequences (the 939 bp promoter in combination with the intron enhancer) are sufficient to reproduce most, if not all, properties of the endogenous *Flk-1* expression.

### **Example 10: The 5'-UTR of the *Flk-1* gene is required for expression of *Flk-1* in the yolk sac vasculature**

In *Flk-1/lacZ* knock-in embryos, the *lacZ* gene is under control of all endogenous regulatory elements except for the regions from bp +137 to bp +299 in the 5'-UTR

and approximately the first 600 bp of the first intron (Shalaby, 1995). It has been shown in accordance with the present invention that the intron sequences deleted in the knock-in construct created by Shalaby (1995) are not required to generate the strong and complete endothelial-specific reporter gene expression which is mediated by the *Flk-1* regulatory sequences described in this study (Fig. 11B and Table III). However, since the complete *Flk-1* 5'-UTR is present in the reporter gene construct which directs the most complete vascular-specific *lacZ* expression (-640 bp/+299 bp//+1677 bp/+3947 bp; Fig. 11B and Table III), it allows to study the consequences of a partial 5'-UTR deletion on *Flk-1* expression in vivo: Genomic DNA was prepared from unstained embryos or yolk sacs. Genotyping was performed by PCR analysis using the primer pairs -258fw/LacRev or LacZP1/LacZP2. Primers for PCR analysis were: -258fw: 5'-ATGGTACCCAGGTTGCTGGGGGCAG-3' (SEQ ID NO: 12); LacRev: 5'-TGGTGCCGGAACCCAGGCAAA-3' (SEQ ID NO: 13); LacZP1: 5'-ATCCTCTGCATGGTCAGGTC-3' (SEQ ID NO: 14); LacZP2: 5'-CGTGGCCTGATTATTCC-3' (SEQ ID NO: 15). The complete vascular staining of the *Flk-1/lacZ* knock-in embryos at E11.5 indicates that the 5'-UTR is not essential for vascular expression in the embryo proper. However, the yolk sac staining pattern of *Flk-1/lacZ* knock-in embryos and of transgenic mice from this study that harbored constructs containing the complete 5'-UTR were markedly different (Fig. 14A-C). The uniform vascular *lacZ* expression in the transgenic yolk sacs from this study (Fig. 14A, B) was absent in small vessels of the yolk sacs of the knock-in embryos (Fig. 14C), in which only large yolk sac vessels were stained. In addition, it was found that replacement of the entire *Flk-1* promoter including the 5'-UTR by the *tk* promoter in the present transgenic construct (Table III) leads to a similar *lacZ* expression pattern in the yolk sacs as that described in the yolk sacs of the *Flk-1/lacZ* knock-in embryos. Thus, the 5'-UTR might be involved in specifying *Flk-1* expression in a subset of endothelial cells.

**Example 11: The role of HIF-2 $\alpha$  in *Flk-1* regulation**

The *Flk-1* promoter (-640 bp/+299 bp) confers endothelium-specific expression to the firefly luciferase reporter gene in transfected bovine aortic endothelial cells (Rönicke, 1996) and provides for a strong reporter gene transcription in vivo; see Examples 6 to 10. This suggests that transcription factors that are specifically expressed in endothelial cells activate the *Flk-1* promoter in a cell-type specific manner.

The basic helix-loop-helix PAS-domain transcription factor, HIF-2 $\alpha$  (also known as HLF, HRF or EPAS1), is expressed in endothelial cells during mouse embryonic development (Ema, Proc. Natl. Acad. Sci. USA 94, 4273-4278, 1997; Flamme, Mech. Dev. 63 (1997), 51-60; Tian, Genes Dev. 11 (1997), 72-82) and is thus a candidate to regulate *Flk-1* expression. HIF-2 $\alpha$  has previously been shown to stimulate both the expression of VEGF (Ema, 1997) and *Tie2* (Tian, 1997). To determine if HIF-2 $\alpha$  might be involved in the regulation of *Flk-1* gene expression, A293 cells were co-transfected with a luciferase reporter gene construct containing *Flk-1* promoter sequences (-640 bp to +299 bp) and an eukaryotic expression vector that contained the mouse HIF-2 $\alpha$  cDNA. Mouse HIF-2 $\alpha$  and HIF-1 $\alpha$  cDNAs were obtained from a mouse brain capillary endothelial cell cDNA library (Schnürch, Development 119 (1993), 957-968) with a 300 bp BamHI/NcoI fragment spanning the 5'UTR of HIF-1 $\alpha$ . Positive phages were rescreened, and inserts were amplified by PCR using oligonucleotides HIF Start: (5'-GGGAATTCACCATG AGTTCTGAACGTCGAAAAG-3'; SEQ ID NO: 18) and HIF Flag Stop: (5'-AAGCGGCCGCTCATTTATCGTCATCGTCCTTGTGAATCGTTAACTTGATCCAAG CTCTG-3'; SEQ ID NO: 19). The PCR product was digested with Eco RI and NotI and subcloned in the EcoRI and NotI restriction sites of pcDNA3 expression vector. The murine HIF-2 $\alpha$  cDNA was obtained as described (Flamme, 1997). The phage insert was amplified by PCR using oligonucleotides HRF START (5'-GGGAATTCACCACAATGACAGCTGACAAGGAG-3'; SEQ ID NO: 20) HRF rev (5'-AAGCGGCCGCTCATTTATCGTCATCGTCCTTGTGAATCGTTGGTGGCCTGTCCA GAGCTCTGAG-3'; SEQ ID NO: 21) and PCR product was digested with EcoRI/NotI

and cloned into the EcoRI and NotI sites of pcDNA3. The sequence encoding the FLAG epitope was included in the reverse oligonucleotide primer. HIF-2 $\alpha$  and HIF-1 $\alpha$  expression plasmids were constructed by inserting the FLAG-tagged cDNAs into the EcoRI and NotI sites of pcDNA3 (Invitrogen). For co-transfection assays, A293 cells were split 1:2 into 35 mm dishes and transfected 18 hours later with 4  $\mu$ g of DNA (2  $\mu$ g of *Flk-1* promoter-driven luciferase plasmid, 1  $\mu$ g of CMV promoter-driven  $\beta$ -

galactosidase expression plasmid, and 1  $\mu$ g of the HIF-2 $\alpha$  or HIF-1 $\alpha$  expression plasmids, or pBluescript SKII and pcDNA3 as a control) using a transfection kit (MBS, Stratagene). After 20 hours, reporter gene activity measurements were performed using the Dual Light Kit (Tropix, Bedford, MA). The luciferase activity of each extract was normalized to the respective  $\beta$ -galactosidase activity. Endogenous background levels of both enzyme activities were measured using extracts from mock-transfected cells and were subtracted. The normalized luciferase activity of the control transfection was arbitrarily set to 1. Each value represents the average of at least six experiments.

In comparison to cells transfected with the luciferase reporter construct alone, co-transfection of the HIF-2 $\alpha$  construct increased reporter gene activity approximately 15-fold (Fig. 15). In contrast, HIF-1 $\alpha$ , a close relative of HIF-2 which stimulates the hypoxia-induced transcription of the *VEGF* gene, failed to stimulate the reporter construct (Fig. 15). These results suggest that HIF-2 $\alpha$  regulates the expression of the *Flk-1* gene.

### Summary

The mouse *Flk-1* receptor is crucial for the differentiation of the hemangioblastic lineage and during embryonic vascular development (Risau, Annu. Rev. Cell Dev. Biol. 11 (1995), 73-91; Shalaby, 1995; Risau, 1997). Moreover, *Flk-1* plays a central role in the regulation of neovascularization in a wide variety of tumors (Plate, Brain Pathol. 4 (1994), 207-218; Ferrara, 1996). To elucidate the basis of its endothelial expression, regulatory sequences of the murine *Flk-1* gene have been isolated and



characterized that confer endothelium-specific reporter gene expression in transgenic mouse embryos. Transgene expression driven by these sequences was strong, specific, and highly reproducible. Most importantly, it has been demonstrated that the isolated sequences were active in early stage vascular development and may thus represent a clue towards the identification of the molecular mechanisms involved in hemangioblast differentiation and vasculogenesis. Moreover, transgene expression persists until shortly after birth and is downregulated in adult animals, as it was described for the endogenous *Flk-1* gene (Millauer, 1993; Kremer, 1997).

Endothelium-specific expression in almost all transgenic mouse embryos tested was mediated by a 939 bp fragment of the promoter region in combination with a fragment of the first intron. 5'-flanking fragments up to -5.5 kbp alone were not sufficient to confer a reproducible endothelium-specific transgene expression. Reproducible endothelium-specific expression was therefore dependent on sequences from the first intron. These sequences also activated the heterologous *tk* promoter specifically in endothelial cells in vivo, and were active in an orientation independent manner. Thus, they fulfill the criteria for an autonomous tissue specific enhancer.

As demonstrated in Example 8, the intron sequences that were sufficient for endothelium-specific expression were contained in a 510 bp fragment. Several potential binding sites for known transcription factors could be identified therein (see Figure 12), including consensus binding sites for c-ets1, PEA3 (an Ets-like transcription factor), GATA transcription factors, and Scf/Tal-1. The c-ets1 transcription factor was proposed to be involved in the early differentiation of endothelial cells from their precursors (Pardanaud, Cell Adhesion and Communication 1 (1993), 151-160). In addition, c-ets1 is expressed in endothelial cells during tumor vascularization and other forms of angiogenesis in humans (Wernert, Am. J. Pathol. 140 (1992), 119-127). Proteins of the Ets family can activate transcription through a PEA3 motif (Wernert, 1992). Transcription factors of the GATA family are involved in the transcription of genes that are expressed in the hematopoietic and endothelial lineages, such as *von Willebrand factor* (Jahroudi, Mol. Cell. Biol. 14 (1994), 999-1008). Unlike the hematopoietic-transcription factor

GATA-1, GATA-2 is expressed in both the endothelial and hematopoietic lineages (Elefanty, Blood 90 (1997), 1435-1447). Scl/Tal-1 has recently been implicated in the regulation of *Flk-1* expression in Zebrafish (Liao, Genes Dev. 12 (1998), 621-626). The presence of two potential Scl/Tal-1 binding sites in the murine *Flk-1* intron enhancer suggests that Scl/Tal-1 might regulate *Flk-1* expression in mice. However, no direct effect of Scl/Tal-1 on *Flk-1* expression has been observed so far in mice, although Scl-null mice have vascular defects (Visvader, Genes Dev. 12 (1998), 473-479).

Recently, analyses of the regulatory elements of other endothelium-specific genes such as *von Willebrand factor* (Aird, Proc. Natl. Acad. Sci. USA. 92 (1995), 4567-4571), *c-ets-1* (Jorcyk, Cell. Mol. Biol. (Noisy-le-grand) 43 (1997), 211-225) or the endothelial receptors, *Tie1* (Korhonen, Blood 86 (1995), 1828-1835) and *Tie2* (Schlaeger, Development. 121 (1995), 1089-1098; Schlaeger, 1997) have been reported. The most uniform expression pattern reported was conferred by regulatory elements of the *Tie2* gene. However, in contrast to *Flk-1*, expression of *Tie2* and of reporter genes driven by *Tie2* regulatory sequences is not downregulated in adult animals. Such as in the *Flk-1* gene, the first introns of the *Tie2* gene and of the *Ets-1* gene are involved in endothelium-specific expression. Similar to the *Flk-1* intron enhancer, the first intron of the *Tie2* gene also contains an autonomous endothelial specific enhancer. A major difference between the structural organisation of the regulatory elements of the *Flk-1* gene and the *Tie2* gene is, however, that the *Tie2* promoter by itself is active in certain embryonic blood vessels (Schlaeger, 1995). At least during the developmental stages analyzed (i.e. E10.5 and E11.5) an autonomous function of the *Flk-1* promoter was not observed. The intronic 303 bp *Tie2* core enhancer also contains potential binding sites for transcription factors of the Ets and GATA families (Schlaeger, 1997), and c-ets1 or PEA3 binding sites are present in the promoters of *Tie1*, *Tie2* and *Flt-1* (Korhonen, 1995; Schlaeger, 1995; Wakiya, J. Biol. Chem. 271 (1996), 30823-30828).

Analysis of *Flk-1/lacZ* knock-in mouse embryos that express the *lacZ* gene from the endogenous *Flk-1* locus has previously shown that the *lacZ* reporter gene is expressed ubiquitously in the developing intra-embryonic vasculature and the yolk

sac of E7.5 embryos (Shalaby, 1995). However, in accordance with the present invention it was found that a fragment of the 5' UTR that is deleted in the knock-in construct is required for reporter gene expression in the yolk sac vasculature during later stages of embryonic development. Based upon transient transfection analyses in bovine aortic endothelial cells, the *Flk-1* 5'-UTR has been shown to contain a positive acting, endothelial cell-specific element between nucleotides +136 and +299 (Rönicke, 1996). The complete vascular staining of the *Flk-1/lacZ* knock-in embryo proper at E11.5 demonstrates that the 5'-UTR is not essential for intraembryonic vascular expression at this developmental stage.

The involvement of HIF-2 $\alpha$  in the regulation of *Flk-1* expression further emphasizes the role of basic helix-loop-helix/PAS-domain transcription factors in the regulation of components of the VEGF signal transduction system and of vascular development. The upregulation of VEGF in response to hypoxia is generally thought to be mediated by HIF-1. Moreover, mouse embryos lacking functional genes for HIF-1 $\alpha$  or ARNT show defects in vascular development, perhaps due to reduced VEGF levels (Maltepe, Nature 386 (1997), 403-407; Iyer, Genes Dev. 12 (1998), 149-162). This observation indicates that the physiological relevance of these transcription factors is not restricted to adaptation to hypoxia, but extends to the regulation of normal vascular development. HIF-2 $\alpha$  is expressed in various tissues, including the developing endothelium of several organs, for example in the brain (Flamme, 1997). It seems therefore likely that HIF-2 $\alpha$  is involved in the regulation of *Flk-1* expression in blood vessels that co-express HIF-2 $\alpha$  and *Flk-1*. Interestingly, HIF-2 $\alpha$  is also expressed in tissues that express the *Flk-1* receptor ligand, VEGF, and has been shown to stimulate VEGF expression (Ema, 1997). Taken together, these observations support the hypothesis that HIF-2 $\alpha$  is both an intrinsic and extrinsic regulator of blood vessel growth and function (Flamme, 1997), by stimulating both receptor and ligand expression. The expression of VEGF and *Flk-1* shows a remarkable coordinate temporal pattern both in development and in tumors. For example, VEGF and *Flk-1* are expressed transiently in the developing mouse brain, and are largely down-regulated in the adult, but reactivated in brain tumors (Plate, 1994). In hemangioblastomas of the brain, which are highly vascularized tumors,

both the *VEGF* and *Flk-1* expression are highly up-regulated, and this correlates with the up-regulation of *HIF-2 $\alpha$*  expression in the stromal cells of this tumor type. Whether *HIF-2 $\alpha$*  contributes to the remarkably coordinated expression of *VEGF* and *Flk-1* in other tumor types, remains to be established, since for example in glioblastomas - another cerebral tumor - up-regulation of *VEGF* is due to hypoxia, and *HIF-2 $\alpha$*  is inducible by hypoxia. Unlike the expression of *VEGF* and *Flt-1*, *Flk-1* expression is not directly stimulated by hypoxia (Gerber, J. Biol. Chem. 272 (1997), 23659-23; Kremer, 1997). Thus, the primary function of *HIF-2* in the regulation of *Flk-1* expression does not appear to be related to the hypoxia response.

Among the endothelial RTK identified thus far, *Flk-1* is the only receptor whose function is required for the determination of the endothelial lineage. Therefore, the *Flk-1* gene represents the ideal candidate for studying the transcriptional regulatory mechanisms that are active during the emergence of the endothelial lineage. The observation that the isolated regulatory elements of the *Flk-1* gene are active in early stage vascular development are of great importance for this objective. Knowledge of the *Flk-1* gene regulatory sequences is also of great potential relevance in the therapy of certain diseases. The *Flk-1* receptor has been demonstrated to be a key regulator of angiogenesis in various diseases, including cancer (Plate, 1994). Therefore, the study of the regulatory elements involved in the upregulation of *Flk-1* expression in the tumor endothelium appears to be particularly relevant for studying the mechanisms of tumor angiogenesis. Further studies will unravel whether the same regulatory elements of the *Flk-1* gene that confer endothelium-specific expression in mouse embryos are also active in the tumor vasculature. *Flk-1* gene regulatory elements active in the tumor vasculature may provide information about the signaling pathways that can be targeted for anti-angiogenic tumor therapy. Finally, the *Flk-1* gene regulatory elements will be useful for targeting expression of genes to the vasculature. An attractive possibility is the expression of suicide genes (Ozaki, Hum. Gene Ther. 7 (1996), 1483-1490) under the control of these elements. The use of the *Flk-1* gene regulatory elements in combination with, e.g., the Cre/loxP system may provide a powerful tool for specifically inactivating genes in the developing vasculature or in tumor endothelium.

The present invention is not to be limited in scope by its specific embodiments described which are intended as single illustrations of individual aspects of the invention and any DNA molecules, or vectors which are functionally equivalent are within the scope of the invention. Indeed, various modifications of the invention in addition to those shown and described therein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Said modifications intended to fall within the scope of the appended claims.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: Max-Planck-Gesellschaft zur Foerderung der  
Wissenschaften e. V.  
(B) STREET: none  
(C) CITY: Berlin  
(D) STATE: none  
(E) COUNTRY: Germany  
(F) POSTAL CODE (ZIP): none

(ii) TITLE OF INVENTION: Regulatory Sequences Capable Of Conferring  
Expression Of A Heterologous DNA Sequence In Endothelial  
Cells In Vivo And Uses Thereof

(iii) NUMBER OF SEQUENCES: 21

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12845 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TCTAGAATAT AGAAGATAAG TTGCGTACA ATTCAGTCAT TTGAAGACCT GATAAGCTTT	60
AAGAAGGAAG ATGGGTTACA CATTGGGAAA TGGTTGCAAT CTGCACATGG CAGAGGCAAG	120
AGATGCAAAT ACATTTTCTT ACATACTCCA TACAAATCTT ACAAGACTGT TTTTCTTTCT	180
CATTTAAAT AAGAAGACCT GCCAGTCTTC CCCTTATTAC TAATTACAGT CACTCTGTAT	240
CTTTGTTGAC ATTGGATAGT TTTACATACT TCAACAGGCT GGTGTCATT AAGTTGTGGT	300
GGGTGGGCAC CAGAGACACG TGATTTCAGAG TGGGAGGAGA TGCAGGAGAA ACGAGGCACA	360

GCAGAAGCAG AAGCGAGGAA AAACACTCTC AACGTTACTA ACACATCGAG AGGTTCCGCA	420
CACTAGCAAT ACGGGCTGAA TCTGACCTAA TCTCTGCTGT TGAAAAATTT GCCTAGCCGC	480
ACACTAGCAA TACGGGCTGA ATCTGACCTA ATCTCTGCTG TTGAAAAATTT TGCCTAGCCT	540
GTCACACAAG TGCTGAGCAT ACAGAAAAAG GAGAGTAATT CTCTGGTTCT TGTACTAACC	600
AAATAGTCTA TATCAAATTG CCTAAGATAA TGTATACATT TAGTACATGA CTGGTTATAC	660
CTATTCTATA TGACTATTAT TTAATGTGA ATTTACAAGT GAGCATATGA AGTCCATTTT	720
ACATGGCTAG TACATATAAC TTTTAAAAAG TTGGACATAG TTATATTTTT CCATTTATTT	780
ATTTACTTTA TATCCTGATC ACAGACCCCC CCCTCCTCTG GATTAAGTCT CTCCACTGCT	840
TCTTACCCCT CCCCATCTCT CTTTCACCTC TGAGAAGGGG GGATACCTCC TGTCTTATCT	900
GGTTTCAGTG GGAGAAGGAT GTATCCTAAC ACATATAATT TTTAATATCC TGAGTTTTTC	960
TTTCATACAC CTACTTATT CTATTCATTT TTCAGGAAGG CATGTTTAAT GTTTTTTTTT	1020
TAATTTTATG TGTACGAGTG TTTTGCTAC ACAGTCATAG TGCATCGCAT ACATTTTTCG	1080
TGCCCGTAGA GATCAGAAGG GAGCATTGGG TTCCCTAGGA CTGGAGGCAT GAACCACCTT	1140
GTGGGTGCAG AGAACTGAGC CTGGGTCATC TCAAAGCATC AGGTTCTTCT TGAGTCATCT	1200
CACTTGCCAC TTCTCCCATI TACTGATTTT ATCTGTGTGC AGACATTCAT GGCCAGTCC	1260
ACAGGTGGAA GTCAGGGACA ACCTATAGGA GTCAGTCCTC TCCTTCTACC GTGTGAGTCC	1320
CTGGCCTCAA ACTCAGGTG TCGGGCTTCA TAGCAAGAGC TTCTATTGTG TGAGCCATCT	1380
TGCTAGCCCC ACCCCATACT ATCTTTATAA TATCTGTTTA ATTAAGACAT TCATAATGAA	1440
TTTTATTAAAC ATTATCGTT ATCCCTTTA CCAATTTTAC TAGTATTAA TTGCCACCCC	1500
TTTAAATTTA ATTACTTCCT TGGCTGGGTT TTACAGGAGA GTTCCAGGAA GCTAGATGGA	1560
GAGATGGCTC AACAGTTTAG AGCAACGGCT GTTCTTGCAg AGGACCTAGG TTCAAGTCTT	1620
GGCACTCAGA GGTGGCTCAC AATCATCTGT GACTTCAGTT CCAGGGGATC TGAAGAATTC	1680
TTCTGGGCTC CATGGGCATC AACTACACAC TTGGTTCATA GACATACATG CCAGCAAATG	1740
ATTGATCCAT ACATATGAAA TAAACCATAA ACAGAAAAAA AAAAGGAAGG TGAGGGAAGG	1800
AAAAAAAGTT TAAAAAAGG AAAGGAAGGA AGGAAGGGAN NNNNNNNNNN NNNNNNNNNN	1860
NNNNNNNNNN NNNNNNNNNN NNNNTCTCTC CATACTGAAA GATGTCCACA ATGACTAAGG	1920
GAATTTTTTT TAAAGACAA GCACAACGTT TTCTAGGGAT CAAACTCTAT TTGTGAGGAA	1980
GACTGGTGGT TTGAAGATTA CATAGCAGAG TTACATCTAA CATGAGCGTG TTTCCCTGG	2040

ATGGAAGGAG TCTGATAACT TGTCTTCTT TCTTAGTTAG CATCTCAGAG TCCCCCGCT	2100
CCCTTAACAT CCTTTTGTCA CACCATCTTT TTAGGAAAA GGATCATTTA TGGGGATGTA	2160
GTGATTGTGA CAAGAATGTC CCCTGTGGGC TCAGATATTT GAATACTTAG TTCCCAGTTG	2220
GGGGAGCTTT TGTAGGGAGG TTGGGAGGCA CAGCCTGGCA GGAGGAAGCA TGCTAGCAGC	2280
TTTGAGACTA TAAACCTCA TCTACTACCT TGTCTCTTT CTGCATTGTG CTGTGTCTGA	2340
CACTGTGAGA TTCCTGTCTC CGATGCCATG CCTGCCCGCC ATGATAGACT CCTAGCCCTC	2400
TGGAAGGTA ACCTCAGTGA ACTCTCTTCT ATAAGTTTCT TTGCTCCTGG TGTTTTATCA	2460
CTGAACCGGA AAAGCTTGCA GGGAGGTAGG AGGCAGCCTG TGCGTGTGAT TCAATGCACC	2520
TGGCCTTATC CTCGGATGAG ATCGGTCACC AGTCAAAAC TGTGAGCTTG AAGGTCTTGG	2580
GTGCTTAACA TCTATTTTAA CAATCTTAT TTAGCAACTT AGAAGTGTGA AATATTGGAA	2640
AGTACTTTAA ACCTTCTAAA CTCCTCCTC CACACTATGA GAATGTTACA TTTTCTATTC	2700
AGTTATTTTT GAGCAGTAAA CAGATGAATC AAGGAATATG CCCATCAGT CAAGAGTGCT	2760
CCTAAATGGA CTGTCTGTT ATTCAATTAC AGTGTGGCCC CTTGACTTTC ATCGGCACCTC	2820
CTAGCAGAAA ACAAATCCG CCAGATGGAG CTGGAGAGAT GGCTCAGCTG TTAAGAATAC	2880
TTATCCCTAC ACAGGCCCTG GAGCCAGTTC CCAGCACCCA CACGGTGGCT CACAACCATC	2940
TGTAAGTCCA GTTCTAGGAG ACCCGACTCC CTCTTCTGTC TGAAAACACC AGGCACGCGT	3000
GCGGTCTACA TACAAACATG AAAGCAAAAT ACACACATTA CATAAATAAA TCTTAAAAAA	3060
TGATTGCGGG TGGGGGAAGG AAAAAAAGG ATGTTAGAAA ATCGATGTAA CTGTTTTTTC	3120
CTTTTGACAA GATCTAAGTT AGGGAAGGAG AACATTCTCT TACCATCGAA AATAATTGTT	3180
TTCATTGCCC CCAAGTCTGC TAATAGAGCT TGCTACCTTC ATGGCTGTGC TAAGGATGAG	3240
GCAAAGATGG ACTTCAGCTT TCAGACTGTG TCTGCTCAA TGTTGGCTAC TCCTGTTTTC	3300
TGACCCCTT CTCTGTGTGA ATGTGGACTT TCAATTAAT TCCCTGCATC TTTTACATAT	3360
TTGATTTAAA AAATATTTTA TTTTATGTAA TTGTATGTAT ATGCATGTCA ATAAGCATAT	3420
GTGTGTGTGT TTCCATGGAA ACCAAGGCAA CAGATTTTCC AGAGCTGTAG AAATGGGCTG	3480
TGAGACGCCC ACTGTGGGTG TTCGGAACCA AACTCGGTC CTGTGGAAG ACAGCGAGCA	3540
CCCATAATGC AGAGGTATCT CTCAGATTTT ACTTTAAAT TTCAATTTTC TTTTTTTTTT	3600
TTAAGTTCC AAGTAACAT AGGAAAGTAC ATGGGTATAT AGATCCCCAG TACCAAGATT	3660



CTTCCTTTGC AGGTAGCACA ACTTGGTTTG TTTCACATAA AGAATGGAAA GTCATTAAAA	3720
CACTCATCAC ACTGTAAAGT AGAATTGAAC TCTGACAGAA CAAGCGAAGT GAGTCTGACT	3780
TCCAGGTAAC TGAGCCTTCT TTTCTCCTA AAGACACAAG CCATACACAG AGTAAAAATA	3840
ACTTGGGCAT GGTGAGAAGG AAACACGCA GGAGGGCTAG CCAAGTCTGA GAGTCGTGAG	3900
TGTGCTCGGT TTATAAACGG AGCCACCTT GCCAGCGAGG TAGTCACATG CTCTGCTAAA	3960
CAGAAACTTA AGAAACACT TACACGAAGC AAACATGGGG AAGTGCCTAG CAAGCATGTG	4020
ACTGACTGGT GGCAATGACC GAAACACAG CAGCCACTAG AAAAGGAAGG GTAGTGCGCC	4080
ACACTGTAGT TGTGAAAATG AACTTATTCA TTTATTTTGA AAAACGTGTA AGAAGCAAAG	4140
ATGTCTTCTT TCCCACCTAC CTTTGGCGCA GCGGAGCACT TCCTGGAATT TATAAAGTGC	4200
GATCTTTCTG GGGACTTCTC ATAACATTTC CTACTGCTCA TCTATGCTG TGTCAAATAG	4260
AGAATGCTCT TGAACAAGTG TGTGTGTGTG TGTGTGTGCG CGCGCACGCG CACTCACTCC	4320
TGCTCTGTTG AGTCCAGTT TTGATGTTCC CGCCAGAGGT ATATTTGAGT ATCATTTCTC	4380
AAGAGCTTCA GCTGGGAGAC ACTGCCTCTT ACTGGCCTGA AGTCACTAG CTGATTTCATC	4440
TCCGTTTGGG CTGGCGCGCC TTGGGGATCC TCCTATCTCT CCTTCCCCAG TGCTGGGATA	4500
ACAAGGTTGG CACCACATGA GCCTTTTAAA ATGTGAGTTT GGAAGCTCAA ACGCAGGTTT	4560
TCATGCTTGC ACTGAACTT CACAAGCTGA ACCGCTCTCC TCTCCTTCCC TCTCTTTTTT	4620
CCTTTTCTTC TTCTTTTAA AACACATCT TGTCTTTAAA AAAAAAAAAA GGCCCCAAAC	4680
AAGTGTAAG TATTTCCCTA TGTGTGTGGA GGGAGGGAGT ATAGGAGGCT GATTTCAGTG	4740
AGATCTGTGT AAATTTGGGT GCCATAGCCA ATCAAAGACG CATCGTTTCC TCTAAGAATT	4800
CTAAATGGGG CGATTACCAC GGGCCTGCAG GTTCTGGTTT GTATTAGAGG AGACACTGTC	4860
TTCTTAAGTA AAACATAGAA GGGGAAGTGT CCAGAATTGT AATAAGGCT TCGAGAGAAG	4920
CCTTGTCTGG CCACCGGGAT GGAGAAGACC TACCTTCGCC TATCCAGGAT CCATCGTCCC	4980
TCCCTCTACC CAGATCTGAC AGCCCTCCTT GGCTCTTTTG CTGAGGTTTG TTTAGTTTG	5040
TTTACTCTC TGCAAGAGAA GTTTCTTAA ACATTCTACC CTGTTACAA GTAAATACAC	5100
CTCTTAGCTA AGAGGCCACA CACCCAGGGG GAACACCGAT AAAAAGAACA AGCCAGAACC	5160
TTCAGAACGC TGTCGATAGG TACACCAAGC AGCCTTCATA CGGAGTTTTC ATTCTGTAGG	5220
AGCTGAATAT ACAACAAGC TAAATGTGAG CAGACCAGGC ATGCCTCTGC TAAATGAGGA	5280
TGCCACACC AAACATGCCC AAGATCTTCA AGTATAATTT TATTATATAG ATTCGCTATG	5340

TGTTGACATG TTTTATAGT GAACCTGGAT TTTACAAACC CTCCTGGTTT GCCACCTGCT	5400
TCTGGCACCA TACTTGAGGC TTAGGCACGT GATAAAGGAG CATGCCTGTT TCCCCCTTA	5460
TTTTTTTTAA AGAAAAGCAC CATGTTACAT CATTAATCAT GCATATCAGT GTAGTTTAGA	5520
TCCGATGTAG AGACAATAAT CTTATCTCTT TGCTGGCTG AAAGACTGTC CTTTAAACTA	5580
TCATTCTAAA TGCATTGGT TTTTGCCAGG AGTAAAACAT GTCACAAGAT ATTTGTGTGTC	5640
ATTTCCACAG CGTGAAGGA AAGGAATGGA AAGAAAACCA GGGGTGAAGG CTGCTGTTCC	5700
TCCTAGTCG CTACTTGAAG TCTACATAGC TGGGGGGGGG GGGGGGACTG TTCACATGGG	5760
ACCGGTTTCC TCTTGTGTC TACACTGGCG CCTCTGGCAA AAAACTCTCC CTTCTCTTCC	5820
CCCCAAGCAT ATCTTGGCTG AAAGCTCAGC TCTGAAAAGG GGCTGGCCA AAGTTACTGT	5880
AGGGGACCGT GGTATGGAA CTGGGTAAAC AAAAGCACTC TAGCAGCCAC TGGAAAAGGA	5940
CCGGGGGCTC TTCTCTGTGC ATTTGCCCTG GAACCCTGAC CACCGCCAGC TCCTGTCATC	6000
TCCTTGCTAT GGGTTTCTG GACCGACCCA GCCAGGAAGT TCACAACCGA AATGCTTCT	6060
AGGGTAATC AGGTAATTC GGACGATTTA AAGTTGCCAG ATGGACGAGA AAACAGTAGA	6120
GGCGTIGGCA ACCTGGATAA GCGCCTATCT TCTAATTAAA ACATTGAGAC GGGGCGGGGG	6180
ATGCGGTGGC CAAAGCACCA TAAACAAAA CTTCCAAGTA CTGACCAACT CACTGCAAGT	6240
TTGTGCCCCG AGTACATCTA GGTTCAGGGG TTCTTGTCCT CATGCTCCCA ACTGCGGGCG	6300
GATTTTTGGT CCCTTGGGAC TTTCAGTGCA GCGGCGAAGA GAGTTCTGCA CTTGCAGGCT	6360
CCTAATGAGG GCGCAGTGGG CCTCGTGTT CTGCTGATGC TCCCAGGTT GCTGGGGGCA	6420
GCAAGTGCTC CAGAGCCCAT TACTGGCTAC ATTTTACTTC CACCAGAAAC CGAGCTGCGT	6480
CCAGATTTGC TCTCAGATGC GACTTGCCGC CCGGCACAGT TCCGGGGTAG TGGGGGAGTG	6540
GGCGTGGGAA ACCGGGAAAC CCAAACCTGG TATCCAGTGG GGGGCGTGGC CGGACGCAGG	6600
GAGTCCCCAC CCCTCCCGGT AATGACCCCG CCCCCATTGC CTAGTGTTGTA GCCGCGCTC	6660
TCTTTCTGCC CTGAGTCCTC AGGACCCCAA GAGAGTAAGC TGTGTTTCTC TAGATCGCGC	6720
GGACCGCTAC CCGGCAGGAC TGAAAGCCCA GACTGTGTCC CGGACCGGGG ATAACCTGGC	6780
TGACCCGATT CCGCGGACAC CGCTGCAGCC GCGGCTGGAG CCAGGGCGCC GGTGCCCGCC	6840
GCTCTCCCCG GTCTTGCCTC GCGGGGGCGC ATACCGCCTC TGTGACTTCT TTGCGGGCCA	6900
GGGACGGAGA AGGAGTCTGT GCCTGAGAAC TGGGCTCTGT GCCCAGCGCG AGGTGCAGGA	6960

TGGAGAGCAA GGCCTGCTA GCTGTCGCTC TGTGGTTCTG CGTGGAGACC CGAGCCGCCT	7020
CTGTGGGTAA GAAGCCCACT CTITAGTAGT AAGCGGAGA AGTAGGGTGC GGGCGGAGAG	7080
TGGGAATAGA AGAGGACCTA ACTCGTAGAG CTCTAGAGAC CCTCCTCCCT TGGGTGTTCT	7140
TTCACTTACC AATGGGGAAA CTGAGGTTC AAGACTTCT CGAAATGACT CAGCCAGGAT	7200
TCTACTCTCC CCCGGGCATC GGTGGAGCG TGTCTGCGG AGCCCTCACA GCCCTGGCG	7260
CTAGGTAGGC AGGAGTGGAA AGGCGGCCTG AGCCGGGGCA GGAGATGCTC CCACTGSCAG	7320
GAACAGGCGG TCAAACGCTG GGAAGCCAGC TCAAGCCAAG CGGCCCGGCT GGCATCAATC	7380
ACTCGTGCT GTTGCCACC GCCCTAGTGG GGGGCAGGGA ATCCGCCTCT GGCCTCGCTC	7440
CCCTTTAGCT CCAGCGTGA AGCGCACGGA CTATGTGAGG GTAGGTCTCT TCATAGAGCA	7500
ACACTTTCT CCTCAACTT TCTTTGATGC AGAATGCTAT TTTTGCTGGT AGGAGGAAGA	7560
CGCGGGTTTC TCTTCGTGA CAGCTTCTCC AGGTGTATTA AACTAAATAA CTCTCCACTT	7620
ACCGACTCCA AAGCGCTGGT CCTGGGGTAA ACTCTGAAAG TCTCAGAAAC TCTTGAGCTT	7680
GGCACCTAGT TATAGGTCAC TTTTCTTGT TTTAAATGCC CTCTGCTTCA AGGTAGGCC	7740
CACACTCGT CTGGGCTTT TGTGCAATAA TTCCCTTCC CTCCCTTCC CTCCCTTCC	7800
CTCCCTTCC CTCCCTTCC CTCCCTTCC CTCCCTTCC CTCTCTCTT TCTCTCTCTC	7860
TCTTCTCTCT CTATTTCTCT GTCAITTCCT TTTTGAAGCC ACAGTTTGCA GATTTCAT	7920
CTCCACCCAT TGGAGAATGG AGAATCAGGA AAAAGAAGT CAATTCTGCA GAAACATTCC	7980
TGCGCCCTA AGAGAATCGC ATGGCTTAA AGCATTGGCA CTGACATACG GCGCCAAGAT	8040
CGCCTGTCTA GAGCTATTGA GTTTCTCTCA TAATGACTTG GTTCATCAGG CTAGCTCCAC	8100
CACGAGTGCC CTCTTGTTC TGAGAAGGCC GCACCTCTCC CTTTCTGGG AAGAGAAAGA	8160
CAGCCTGGAA CATGTGCTTG CCCTGGGTT CATAGAGAAG CAAGTTGCTT TAAAGCCAG	8220
AGAATTCTA GTGTAGCAGC TTAACAGCGT CCCGTTCTCT GAATAAGATG GAGGTTGCC	8280
TTTTGGAGTG TGTGACTTGC TTAATTGGAT TGGGCTATAA TTGGTGCCAT CCAAGTCTCG	8340
AGACAGAGCC GCTGTGTTT TTCTTCTGG TCTTTGAGCG GGAAGGATAA CAGTGACAAA	8400
ATTAATTAAT GTTGTGTTAT GGATTTGAAC ATAAAAGGGC TTTTATTGTA TAGTAGCATA	8460
TGTACCTCTT CGAGTCAGAA TGAGCTGTCT AAAGAACAGA ACCCAAACCT GCCGATGAAA	8520
ATGAATGAGG TTTAATAAAG GCGATGGATG AGCATTAGTC ACTGATGTAA ATCTCCAGTT	8580
ATTGATAACC TCATTGACTG GATTTGATTG CAGACATGTA TTGGTATGGG GCATCCTTTA	8640

AAGATGAGCA TAGCCAACGT GCCTGCACTC TAAGAGAATC TATGGCTGTA TGTTATTACA	8700
GAGACAGTTG AGAAGCTCTT AGTGGCTCTG GCGTGTAGAT CAGCGGTAGA GCGCTGAGGC	8760
CTCTGCGCTCG CTTCCTGGCA CTGAAGAATA AAGGCCATTT ACTGTGGTGG TGCASTGGGC	8820
GCAGTTTGTG ACGAGTTACT ACTACATTTT CCTCACACAT CTGCTGACT AATGAGTTCA	8880
TCAGATGAGC GTATCCAGTG ATTGTTTGCA GGTAAATGGT TCTCAGTCAT GTTTAGAATC	8940
TACTTATCAA ACAAAATGTT TTCTCATTTT CTGCTTCTTC TCAAACAAAG TAAGATTCCA	9000
TTATTGAAAG GCTTGTTTAA GAGCATTTTA ACTGCTTGCC TATGTTAGGG ACAGTGACTT	9060
ATTTCATATT GACAAATATT ATGCCGATTA ATTGAATATG ACTACCCAGT TCTATAGCTG	9120
TCTCAGGCA GACCAAGAGC ATCTGTGATC CAGTCACTTT AAATGCCATT TAAATGCAT	9180
AATTTGTTGG TCTAGGAATA AACACACTGT AAGTTTAGA ATCAGGCCCC AAACACAAGT	9240
CTTTAACAAAT GCCAACTAGC TTCTGAGATT CATTAATGTC ATTTAATTAC CAATGTTTTA	9300
AAAATATGTC ATTAATTACT AAATCTATAG TTGTAACAGC AACACATGTA CATCTTATTA	9360
AGTTGGGTAT ATTCAGGGTG GCATAGCTGT AGACTATTGC ACATCTGTGT TGGTGAGCCA	9420
GTGGAGAACT GCCTCCTGGC TGTCTCAGA AGGCCACAGT GTCACGGCAT TGGCTATTTG	9480
CCTTGGCTCT TTGCTAATAC TTTATTGACA TGGCCTCATC TTCGTTACG TTCACTTATT	9540
TGCCCAACAA CGTCAATGCC AGCTGAGGCC TTAGGAGTCA TCTGTTCTTA GTCAGTGCGA	9600
ATTAGAAAGC CTGGATGCCT GCCTGCTATT AATTAGTTAT TCTTCTCTTC TGAGACAGAG	9660
TCTCACTGTG TGCCCCAGGC TAGTCTCAA CTGCGGTCC ATTTGTCTCA CTCATCAGAA	9720
TGCTGGGCTT CCAGGTGTGT GCACCACACT AGGTAGCTCG CGTTTAAAG TAAGAGCTGG	9780
AAGATCCTGA TGTCCTTTAC CATGGTGGGC ATGTTACAGS TTAGTTGACT GAAACTAGT	9840
TATCTCGCTG TGTAATGACC TGCAGTGGTA TGTATCTCTC AAGATGCTTT TTTGCATTTT	9900
AATCAGTTAG GTAACAAGTT CTTAAGTCTC CAGCTTGGA TTGGCATGAG CTCAGAGCTT	9960
TGATTATGA GTTGGGACCC CCTAGCTATT GCTCATTAGA CTTACACTAT TTTTAGTTTT	10020
GCTCTGAGTT TATGAATATG CATGTATGCA TGAACCTGGG AGATATTTTT CTTCCTCAAT	10080
TCCTTTTCTT CCATTTTAAAT GTGCTGTCTT TAGAAGCCAC TGCCTCAGT TCTGCAGCTC	10140
AGATACAAA GGAAGCTGCG TACACAGCAT GATAAAGAC AATGGGACGG GGTACAGTG	10200
GCTCCGCTCC CTTCAGGGG TATGGAGAGC AGCTGTAGAG AGATGTCTCC AGGAGTTTTT	10260

CATTAATCAG CAATTAGTC AGATCTGTGC ATCCTATGCT TTACAAGAAA TGTCAGTGGG 10320  
CCTGAGATCA TCAGATGGAG GTTCATCGGG TTTCATATGC CCGTATCCTT TTGTAAGACC 10380  
TTGAAGTTGG CAACGCAGGA AAACAGGAAC TCCACCCCTG TGCCGTGAAT TGCAGAGCTG 10440  
TTGTGTGTGT TTGTGACCAT CTGCCCATTC TTCTGTATT GACAGAGCTT GTGAACCTTA 10500  
ACTGGGACTG GGGCAAAGTC AATCCCACCT TTATACAATG AATTGCTGAA GAGGCTTTTT 10560  
AAAACCTGGA GTGTGCATTG TTTATGGAAG GGCTTTCTTA TTGGATCCAA CTCTTTTCTA 10620  
ATTTGTTTCT AGGTTTGCCCT GGCATTTTTC TCCATCCCC CAAGCTCAGC ACACAGAAAG 10680  
ACATACTGAC AATTTTGGA AATACAACCC TTCAGATTAC TTGCAGGTAA GGATTCCTTT 10740  
TTGAGCCAGC TTTCTATGT GAAAGGACTC ATTGTTTACT GAGGTCACAA CAATTTCCAC 10800  
TATTGCAGAA GTATAATAGT ATTGTTACAA TTGTTTATAA ATCATGAGAC TTCTAAGAAC 10860  
CTATTTAATA ATGAAACAT GGAAGAAATC TTTTCAAACC TTGTACTCT TTTGCTGAGC 10920  
CGTTTTCAAC ATGCACAAC ATATTACACA AATATAACAT ACACAGGAAC ACACATGAAT 10980  
GCATGGGATG ATGTGCCTAA AACTAGCATG TAATTGATAT TCACAATTAT TGATAAATTA 11040  
GTAAAGCAAA GGAATTCTCT ATGAATAGAG CTAAATTTCT ATCCATGTTT AAGTCACCCA 11100  
GAATGGCTTC TGGACATTTT TTTTTTAGC TGTCTTCTAC AAGTGAAATT CTGCCGTGAT 11160  
TAGCAATTTA ATAICTAGCC AATAATATTC CTGACCATAT GTCTGTGTCA GACCATGACC 11220  
TTCATAATCT GGCTTGATGT TCTGGGCTTC TTTCCCTCTT GCCAGCAAGA TGTCACGGTG 11280  
TTGATGTCTG ATAACTGAG AAACAGAAGT TTTTCGCAAG AAGAGGACCT TGAATTTTGC 11340  
TTTTCCCTG AGAGACAAGA AAGGAAACTT AGAGGAGGTG TAGCTGGGAG TGTGGTCATT 11400  
CATGAAAGAC CTGTTTGAGC GGCAGTGTGT TTTGCTGGG ACAGTAATGA GCCTAGATCG 11460  
TAGTGCCATC CCAAGAGAGT GCTTGGTGGC AAAAGAGGCC CTAGCAGCTT GTGGCAGTTG 11520  
CCTCATATTT GAAGAATACT AAGAGGTCCC CCGAATAACT CAGGCTAGT GTTGATCATT 11580  
GCATGTGGAG AGAATCCAAG CCTCTATCT AGGGTCTACA AAAGTAACCA ATGCCAGTC 11640  
TTTGGGGGAA AGCAAAACCA GAAAGCGATG ATAGCAGGAC CTGTTTATTT TCAITTAAGTC 11700  
ATGGCATTTC CAGAGACTTT GCTCCCCCTA TTCTCAGACA CAAAGCCAC TTAAGATCTC 11760  
CCTCTGGAGA CTGCTGGGAA CATTTCCTAA GTTCTGAAAA AACCTGGAG TGATTGGGCA 11820  
CAGACGATCC TGTCACTTCA TGTAGTGTCT AAGCTCTTTG GGTGATGACT CAGTGGGTCA 11880  
CAITGTTTTA TTCATATTGA CTACCTTCGG TTTGCTTTCG GGAGAATGGA AGCTATAGAA 11940

GTCTGTTTGG	TGTGGCCCTC	ACAAGGCACT	GTGAGCTTCT	TCTCTCTGTG	TGCTAACTTC	12000
TTACTCTCCC	TTGCTTATAC	CCACATAGGG	ACTCTGGCTT	TGTTGCTGTT	CTTCAATGCT	12060
TCAGATGTGC	CCTGGGTCCT	GTCTGTCCTT	CACACTTACT	GATGCTGCCT	GGAATGCTAT	12120
TCCTCCCAAT	GTGCATAGGG	CCAGCTCGGT	CCTAACTCTC	TCTTTTCTTT	GCCTCTTTTA	12180
TATTTTCCTT	CACAGTATCA	AATCACCACA	GTTTATGCAA	CAAACGTAAA	CTTTAAAAAT	12240
GTCTGTCTCC	TTATATTAGT	GATAGGTTCC	AGAAAGGCAC	TGATTTTTTT	TCTTCCCTGG	12300
TGTACACTGG	GCAACTACTC	TACCACTGAG	CGTGATATCC	TTGGTCCCTT	AAAAGTTATC	12360
CTCTGTCCCT	AATAATGCTT	AGCAATCATA	TTTGCTTAAA	ATATTTATTG	AATGACTGCA	12420
GGAATGAATG	ARTGAATGAG	CTAACAGAAA	ACTCATGACC	ATGTGGGTGA	TTTCCGAAAC	12480
AGAGTGTGAG	ATCTTTGGTG	GCATGTCCTT	GTAGACTGTC	TGCCACCAGT	ATCTATCATC	12540
TTGAAGGTGA	CTATTGAGTA	GTTTATATGC	ATGTGAAAAA	CCAAACCTTC	TATTCTCTTA	12600
CTCATAGCCT	CTCTTAATCA	TAGCCCTGTG	GCATGGAGTG	TACCATTGAT	ATCTTCCTGG	12660
AATACTTTTT	CAGGGGACAG	CGGGACCTGG	ACTGGCTTTG	GCCCAATGCT	CAGCGTGATT	12720
CTGAGGAAAG	GGTATIGGTG	ACTGAATGCG	GCGGTGGTGA	CAGTATCTTC	TGCAAAACAC	12780
TCACCATTCC	CAGGGTGTTT	GGAAATGATA	CTGAGGCCTA	CAAGTGCTCG	TACCGGGACG	12840
TCGAC						12845

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

## (iii) HYPOTHETICAL: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GGGGTACCGA ATTCTAAATG GGGCGATTAC C

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

## (iii) HYPOTHETICAL: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GTGGTACCCA AACACTCAAC ACCACTG

27

## (2) INFORMATION FOR SEQ ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

## (iii) HYPOTHETICAL: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TCGGTACCGA CCCAGCCAGG AAGTTC

26

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

## (iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TTGCTAAGCT TCCTGCACCT CGCGCTGGG

29

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AGGGATCCAC TCTTTAGTAG TAAGGCG

27

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

ACCTCGAGAC TTGGATGGCA C

21

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear



(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGGCTATAAT TGGTGCCATC C

21

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGATGGAGAA AATCGCCAGG C

21

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GTGTGCATTG TTTATGGAAG GG

22

## (2) INFORMATION FOR SEQ ID NO: 11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## ... (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

## (iii) HYPOTHETICAL: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CATAGACATA AACAGTGGAG GC

22

## (2) INFORMATION FOR SEQ ID NO: 12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## ... (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

## (iii) HYPOTHETICAL: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

ATGGGTACCCA GGTGCTGGG GGCAG

25

## (2) INFORMATION FOR SEQ ID NO: 13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## ... (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

## (iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TGGTGCCGGA AACCAGGCAA A

21

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

... (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

ATCCTCTGCA TGGTCAGGTC

20

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

... (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CGTGGCCTGA TTCATTC

18

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GGGAATTCAC CATGATTCT GAACGTCGAA AAG

33

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

AAGCGGCCG TCATTTATCG TCATCGTCTT TGTAATCGTT AACTTGATCC AAAGCTCTG

59

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GGGAATTCAC CACAATGACA GCTGACAAGG AG

32

## (2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 64 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"
- (iii) HYPOTHETICAL: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

AAGCGGCCGC TCATTATCG TCATCGTCCT TGTAATCGTT GGTGGCCTGG TCCAGAGCTC 60  
TGAG 64

## (2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"
- (iii) HYPOTHETICAL: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

CCGGTACCCA AACCCGCCCC AGCGTCTTG 29

## (2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CCGACAAGCT TGGTCGCTCG GTGTTGAGG

30

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197

### Claims

1. A recombinant DNA molecule comprising:
  - (a) at least one first regulatory sequence of an intron of the Vascular Endothelial Growth factor (VEGF) receptor-2 (Flk-1) gene or of an intron of a gene homologous to the Flk-1 gene being capable of conferring expression in endothelial cells in vivo; and
  - (b) operatively linked thereto a heterologous DNA sequence.
2. The recombinant DNA molecule of claim 1, wherein said first regulatory sequence comprises a GATA-binding site, an AP-1 binding site, an SP1 binding site, an NFκB binding site, a STAT binding site, a Scl/Tal-1 binding site, an Ets-1 binding site, a PEA3 consensus sequence or any combination(s) thereof.
3. The recombinant DNA molecule of claim 1 or 2, wherein said first regulatory sequence is selected from the group consisting of
  - (a) DNA sequences comprising a nucleotide sequence as given in SEQ ID NO: 1;
  - (b) DNA sequences comprising a nucleotide sequence of SEQ ID NO: 1 from nucleotide 8260 to nucleotide 10560, from nucleotide 8336 to nucleotide 10608 and/or from nucleotide 10094 to nucleotide 10608;
  - (c) DNA sequences comprising the nucleotide sequence of the human Flk-1-intron;
  - (d) DNA sequences comprising a nucleotide sequence which hybridizes with a nucleotide sequence of (a), (b) or (c) under stringent conditions;
  - (e) DNA sequences comprising a nucleotide sequence which is conserved in the nucleotide sequences of (a), (b) and (c); and
  - (f) DNA sequences comprising a fragment, analogue or derivative of a nucleotide sequence of any one of (a) to (e) capable of conferring expression in endothelial cells.

4. The recombinant DNA molecule of any one of claims 1 to 3, wherein said heterologous DNA sequence is operatively linked to further regulatory sequences.
5. The recombinant DNA molecule of claim 4, wherein said further regulatory sequence is a promoter.
6. The recombinant DNA molecule of claim 4 or 5, wherein said further regulatory sequence is a 3'-untranslated region.
7. The recombinant DNA molecule of claim 5 or 6, wherein said promoter is a promoter of hypoxia inducible genes, genes encoding growth factors or its receptors or glycolytic enzymes.
8. The recombinant DNA molecule of claim 7, wherein said growth factor is VEGF, PDGF or Fibroblast growth factor.
9. The recombinant DNA molecule of any one of claims 5 to 8, wherein said promoter comprises a DNA sequence selected from the group consisting of
  - (a) DNA sequences comprising the nucleotide sequence as given in SEQ ID NO:1 from nucleotide 6036 to nucleotide 6959;
  - (b) DNA sequences comprising the nucleotide sequence of the human Flk-1/KDR promoter;
  - (c) DNA sequences comprising a nucleotide sequence which hybridizes with a nucleotide sequence of (a) or (b) under stringent conditions;
  - (d) DNA sequences comprising a nucleotide sequence which is conserved in the nucleotide sequences of (a) and (b); and
  - (e) DNA sequences comprising a fragment, analogue or derivative of a nucleotide sequence of any one of (a) to (d).
10. The recombinant DNA molecule of any one of claims 1 to 9, wherein at least one of said DNA sequences is of human or murine origin.



11. The recombinant DNA molecule of any one of claims 1 to 10, wherein said heterologous DNA sequence being operatively linked to said regulatory sequences is located 5' to said first regulatory sequence.
12. The recombinant DNA molecule of any one of claims 1 to 11, wherein said heterologous DNA sequence encodes a peptide, protein, antisense RNA, sense RNA and/or ribozyme.
13. The recombinant DNA molecule of claim 12, wherein said protein is selected from the group consisting of Vascular Endothelial Growth Factor (VEGF), Hypoxia Inducible Factors 1 (HIF), HIF-Related Factor (HRF), tissue plasminogen activator, p21 cell cycle inhibitor, nitric oxide synthase, interferon- $\gamma$ , atrial natriuretic polypeptide and monocyte chemotactic proteins.
14. The recombinant DNA molecule of claim 12, wherein said protein is a scorable marker, preferably luciferase, green fluorescent protein or lacZ.
15. The recombinant DNA molecule of claim 12, wherein said antisense RNA or said ribozyme are directed against a gene involved in vasculogenesis and/or angiogenesis and/or tumors of endothelial origin.
16. A nucleic acid molecule of at least 15 nucleotides in length hybridizing specifically with the first regulatory sequence of a recombinant DNA molecule of any one of claims 1 to 15.
17. A vector comprising a recombinant DNA molecule of any one of claims 1 to 15.
18. The vector of claim 17, which is an expression vector and/or a targeting vector.
19. The vector of claim 17 or 18, further comprising a gene capable of expressing HIF-2 $\alpha$ .
20. A cell transformed with a DNA molecule of any one of claims 1 to 15 or the vector of any one of claims 17 to 19.

21. The cell of claim 20, which is a prokaryotic or eukaryotic cell.
22. The cell of claim 20 or 21, which is an endothelial cell.
23. The cell of any one of claims 20 to 22, further comprising a recombinant DNA molecule or vector containing a gene capable of expressing HIF-2 $\alpha$ .
24. A pharmaceutical composition comprising a recombinant DNA molecule of any one of claims 1 to 15, the vector of any one of claims 17 to 19 and/or the nucleic acid molecule of claim 16 and optionally a pharmaceutically acceptable carrier.
25. A diagnostic composition comprising a recombinant DNA molecule of any one of claims 1 to 15, the vector of any one of claims 17 to 19, the cell of any one of claims 20 to 23 and/or the nucleic acid molecule of claim 16, and optionally suitable means for detection.
26. A method for the production of a transgenic non-human animal, comprising introduction of a recombinant DNA molecule of any one of claims 1 to 15 or a vector of any one of claims 17 to 19 into a germ cell, an embryonic cell or an egg cell or a cell derived therefrom.
27. A transgenic non-human animal comprising stably integrated into its genome a recombinant DNA molecule of any one of claims 1 to 15 and/or the vector of any one of claims 17 to 19 or obtained according to the method of claim 26.
28. The method of claim 26 or the transgenic non-human animal of claim 27, wherein said animal is a mouse.
29. A method for the identification of a chemical and/or biological substance capable of suppressing the transcription of a gene in endothelial cells comprising:

- (a) contacting a cell of any one of claims 20 to 23 or the transgenic non-human animal of claim 27 or 28 either of which is capable of expressing the heterologous DNA sequence with a plurality of compounds; and
  - (b) determining those compounds which suppress the expression of said heterologous DNA sequence.
30. A method for the identification of a chemical and/or biological substance capable of activating and/or enhancing the transcription of a gene in endothelial cells comprising:
- (a) contacting a cell of any one of claims 20 to 23 or the transgenic non-human animal of claim 27 or 28 either of which is capable of expressing the heterologous DNA sequence with a plurality of compounds; and
  - (b) determining those compounds which are capable of activating and/or enhancing the expression of said heterologous DNA sequence.
31. Use of a recombinant DNA molecule of any one of claims 1 to 15, the vector of any one of claims 17 to 19, the cell of any one of claims 20 to 23, the pharmaceutical composition of claim 24, the diagnostic composition of claim 25 and/or the transgenic non-human animal of claim 27 or 28 for the identification of a chemical and/or biological substance capable of suppressing or activating and/or enhancing the transcription, expression and/or activity of genes and/or its expression products in endothelial cells.
32. The method of claim 29 or 30 or the use of claim 31, wherein the chemical and/or biological substance is selected from the group consisting of peptides, proteins, nucleic acids, antibodies, small organic compounds, hormones, neurotransmitters, peptidomimics and PNAs.
33. A method for the production of a pharmaceutical composition comprising the steps of the method of claim 29 or 30 and (c) formulating the compound identified in step (b) in a pharmaceutically acceptable form.
34. A method of inhibiting a vascular disease in a subject, comprising contacting an artery of said mammal with the vector of any one of claims 17 to 19, wherein

said heterologous DNA sequence encodes a protein that reduces or prevents the development of the vascular disease.

35. The method of claim 34, wherein said protein reduces proliferation of smooth muscle cells.
36. Use of a recombinant DNA molecule of any one of claims 1 to 15, the vector of any one of claims 17 to 19, the nucleic acid molecule of claim 16 and/or a substance identified by the method of claims 29, 30 or 32 for the preparation of a composition for directing or preventing expression of genes specifically in endothelial cells.
37. Use of a recombinant DNA molecule of any one of claims 1 to 15, the vector of any one of claims 17 to 19, the nucleic acid molecule of claim 16 and/or a substance identified by the method of claims 29, 30 or 32 for the preparation of a pharmaceutical composition for treating, preventing and/or delaying a vascular disease and/or a tumorous disease in a subject.
38. Use of a recombinant DNA molecule of any one of claims 1 to 15, the vector of any one of claims 17 to 19 and/or the nucleic acid molecule of claim 16 for the preparation of a pharmaceutical composition for inducing a vascular disease in a non-human animal or in the transgenic non-human animal of claim 27 or 28.
39. The method of claim 34 or 35 or the use of any one of claims 36 to 38, wherein the vascular disease is atherosclerosis and/or a neuronal disorder.
40. Use of a regulatory sequence as defined in any one of claims 1 to 3 for enhancing and/or directing gene expression in endothelial cells.

-6660 TCTAGAATAT AGAAGATAAG TTTGCGTACA ATTCACTCCT TTGAAGACCT  
GATAAGCTTT AAGAAGGAAG ATGGGTTACA CATTGGGAAA TGGTTGCAAT  
CTGCACATGG CAGAGGCAAG AGATGCAAAT CACATTTCTT ACATACTCCA

-6510 TACAAATCTT ACAAGACTGT TTTTCTTTCT CATTTAAAAA AAGAAGACCT  
GCCAGTCTTC CCCTTATTAC TAATTACAGT CACTCTGTAT CTTTGTGTGAC  
ATTGGATAGT TTTACATACT TCAACAGGCT GGTGTCAATTA AAGTTGTGGT  
GGGTGGGCAC CAGAGACACG TGATTAGAG TGGGAGGAGA TGCAGGAGAA  
ACGAGGCACA GCAGAAGCAG AAGCGAGGAA AAACACTCTC AACGTTACTA  
ACACATCGAG AGGTTCCGCA CACTAGCAAT ACGGGCTGAA TCTGACCTAA  
TCTCTGCTGT TGAAAATTTT GCCTAGCCGC ACACAGCAA TACGGGCTGA  
ATCTGACCTA ATCTCTGCTG TTGAAAATTT TGCCCTAGCCT GTCACACAAG  
TGCTGAGCAT ACAGAAAAAG GAGAGTAATT CTCTGGTTCT TTGACTAACC  
AAATAGTCTA TATCAAATTG CCTAAGATAA TGTATACATT TAGTACATGA

-6010 CTGGTTATAC CTATTTCTATA TGACTATTAT TTAAATGTGA ATTTACAAGT  
GAGCATATGA AGTCCATTTT ACATGGCTAG TACATATAAC TTTTAAAAAG  
TTGGACATAG TTATATTTTT CCATTTATTT ATTTACTTTA TATCCTGATC  
ACAGACCCCC CCTTCCTCTG GATTAACCTCT CTCCACTGCT TCTTACCCCT  
CCCCATCTCT CCTTCACCTC TGAGAAGGGG GGATACCTCC TGTCTTATCT  
GGTTTCAGTG GGAGAAGGAT GTATCCTAAC ACATATAATT TTAAATATCC  
TGAGTTTTTC TTTCATACAC CTTACTTATT CTATTCATTT TTCAGGAAGG  
CATGTTTAAT GTTTTTTTTT TAATTTTATG TGTACGAGTG TTTTGCCTAC  
ACAGTCATAG TGCATCGCAT ACATTTTGTG TGCCCGTAGA GATCAGAAGG  
GAGCATTTGG TTCCCTAGGA CTGGAGGCAT GAACCACCTT GTGGGTGCAG  
AGAACTGAGC CTGGGTCATC TCAAAGCATC AGGTTCTTCT TGAGTCATCT  
CACTTGCCAC TTCTCCCAT TACTGATTTT ATCTGTGTGC AGACATTCAT  
GGCCAGTCC ACAGTGGAA GTCAGGGACA ACCTATAGGA GTCAGTCTCT  
TCCTTCTACC GTGTGAGTCC CTGGCCTCAA ACTCAGGTTG TCGGGCTTCA

Figure 1

2/21

-5010

TAGCAAGAGC TTCTATTTGT TGAGCCATCT TGCTAGCCCC ACCCCATACT  
 ATCTTTTATAA TATCTGTTTA ATTAAGACAT TCATAATGAA TTTTATTTAA  
 ATTCATCGTT ATCCCCTTTA CCAATTTTAC TATGTATTAA TTGCCACCCC  
 TTTAAATTTA ATTACTTCCT TGGCTGGGTT TTACAGGAGA GTTCCAGGAA  
 GCTAGATGGA GAGATGGCTC AACAGTTTAG AGCAACGGCT GTTCTTGCAG  
 AGGACCTAGG TTCAAGTCCT GGCACCTCAGA GGTGGCTCAC AATCATCTGT  
 GACTTCAGTT CCAGGGGATC TGAAGAATTC TTCTGGGCTC CATGGGCATC  
 AACTACACAC TTGGTTTCATA GACATACATG CCAGCAAATG ATTGATCCAT  
 ACATATGAAA TAAACCATAA ACAGAAAAAA AAAAGGAAGG TGAGGGAAGG  
 AAAAAAAGTT TAAAAAAGG AAAGGAAGGA AGGAAGGGAN NNNNNNNNNN  
 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNTCTCTC CATACTGAAA  
 GATGTCCACA ATGACTAAGG GAATTTTTTT TAAAAGACAA GCACAACGTT  
 TTCTAGGGAT CAAACTCTAT TTGTGAGGAA GACTGGTGGT TTGAAGATTA  
 CATAGCAGAG TTACATCTAA CATGAGCGTG TTTCCCTGG ATGGAAGGAG  
 TCTGATAACT TGTCTTTCTT TCTTAGTTAG CATCTCAGAG TCCCCCGCCT  
 CCCTTAACAT CCTTTTTGCA CACCATCTTT TTAGGAAAAA GGATCATTTA  
 TGGGGATGTA GTGATTTGTA CAAGAATGTC CCCTGTGGGC TCAGATATTT  
 GAATACTTAG TTCCCAGTTG GGGGAGCTTT TGTAGGAGG TTGGGAGGCA  
 CAGCCTGGCA GGAGGAAGCA TGCTAGCAGC TTTGAGACTA TAAACCCTCA  
 TCTACTACCT TGTTCCTTTT CTGCATTGTG CTGTGTCTGA CACTGTGAGA  
 TTCTGTCTCC CGATGCCATG CCTGCCCCGCC ATGATAGACT CCTAGCCCTC  
 TGGAAAGGTA ACCTCAGTGA ACTCTCTCTCT ATAAGTTTCT TTGCTCCTGG  
 HindIII (-4200)  
 TGTTTTATCA CTGAAACGGA AAAGCTTGCA GGGAGGTAGG AGGCAGCCTG

Figure 1 continued

BstEII (-4100)

TGGCGTTGAT TCAATGCACC TGGCCTTATC CTCGGATGAG ATCGGTCACC  
AGTCAAAAAC TGTGAGCTTG AAGGCTCTGG GTGCTTAACA TCTATTTTAA  
CAAATCTTAT TTAGCAACTT AGAACTGTGA AATATTGGAA AGCTACTTAA  
-4010 ACCTTCTAAA CTCCTCCTC CACACTATGA GAATGTTACA TTTTCTATTCT  
AGTTATTTTT GAGCAGTAAA CAGATGAATC AAGGAATATG CCCATCACAT  
CAAGAGTGCT CCTAAATGGA CTGCTTGTTT ATTCATTTAC AGTGTGGCCC  
CTTGACTTTC ATCGGCACTC CTAGCAGAAA ACAAAATCCG CCAGATGGAG  
CTGGAGAGAT GGCTCAGCTG TTAAGAATAC TTATCCCTAC ACAGGCCCTG  
GAGCCAGTTC CCAGCACCCA CACGGTGGCT CACAACCATC TGTAATCCA  
GTTCTAGGAG ACCCGACTCC CTCTTCTGTC TGAAAACACC AGGCACCGCT  
GCGGTCTACA TACAAACATG AAAGCAAAAT ACACACATTA CATAAATAAA  
TCTTAAAAAA TGATTCGGGG TGGGGGAAGG AAAAAAAGG ATGTTAGAAA  
ATCGATGTAA CTGTTTTTTC CTTTGCACA GATCTAAGTT AGGGAAGGAG  
AACATTCTCT TACCATCGAA AATAATTGTT TTCATTGCCC CCAAGTCTGC  
TAATAGAGCT TGCTACCTTC ATGGCTGTCG TAAGGATGAG GCAAAGATGG  
ACTTCAGCTT TCAGACTGTG TCTGCTCAA TGTTGGCTAC TCTGTTTTTC  
TGACCCCTT CTCTGGTGCA ATGTGGACTT TCAATTAATT TCCTGCATC  
TTTTACATAT TTGATTTAAA AAATATTTTA TTTTATGTAA TTGTATGTAT  
ATGCATGTCA ATAAGCATAT GTGTGTGTGT TTCCATGGAA ACCAAGGCAA  
CAGATTTTCC AGAGCTGTAG AAATGGGCTG TGAGACGCCC ACTGTGGGTG  
TTCGGAACCA AACTCGGGTC CTGTGGAAG ACAGCGAGCA CCCATAATGC  
AGAGGTATCT CTCAGATTTT ACTTTAAAAT TTCAATTTTC TTTTTTTTTT  
TTAAGTTTCC AAGTAACTAT AGGAAAGTAC ATGGGTATAT AGATCCCCAG  
-3010 TACCAAGATT CTCCTTTG CAGGTAGCACA ACTTGGTTTG TTTACATAA  
AGAATGGAAA GTCATTAAAA CACTCATCAC ACTGTAAAGT AGAATTGAAC  
TCTGACAGAA CAAGCGAAGT GAGTCTGACT TCCAGGTAAC TGACCTTCT

Figure 1 continued

TTTCCTCCTA AAGACACAAG CCATACACAG AGTAAATAA ACTTGGGCAT  
GGTGAGAAGG AAACAACGCA GGAGGGCTAG CCAAGTCTGA GAGTCGTGAG  
TGTGCTCGGT TTATAAACGG AGCCACCTT GCCAGCGAGG TAGTCACATG  
CTCTGCTAAA CAGAACTTA AGAAAACACT TACACGAAGC AAACATGGGG  
AAGTGCCATG CAAGCATGTG ACTGACTGGT GGCAATGACC GAAACCACAG  
CAGCCACTAG AAAAGGAAGG GTAGTGCGCC AACTGTAGT TGTGAAAATG  
AACTTATTCA TTTATTTTGA AAAACGTGTA AGAAGCAAAG ATGTCTTCTT  
TCCCACCTAC CTTTGC GGCA GGCAGCACT TCCTGGAATT TATAAAGTGC  
GATCTTTCTG GGGACTTCTC ATAACATTTC CTACTGCTCA TCTATGTCTG  
TGTCAAATAG AGAATGCTCT TGAACAAGTG TGTGTGTGTG TGTGTGTGCG  
CGCGCACGCG CACTCACTCC TGCTCTGTTG AGGTCCAGTT TTGATGGTCC  
CGCCAGAGGT ATATTTGAGT ATCATTTCTC AAGAGCTTCA GCTGGGAGAC  
ACTGCCCTCTT ACTGGCCTGA AGGTCAC TAG CTGATTCTCC TCCGTTTGGG  
CTGGCGCGCC TTGGGGATCC TCCTATCTCT CCTTCCCCAG TGCTGGGATA  
ACAAGGTGCG CACCACATGA GCCTTTTAA ATGTGAGTTT GGAAGCTCAA  
ACGCAGGTTT TCATGCTTGC ACTGAAACTT CACAAGCTGA ACCGTCTCCC  
TCTCCTTCCC TCTCTTTTTT CCTTTTCTTC TTCCTTTTTA AAACACATCT  
-2010 TGTCTTTTAA AAAAAAAAAA GGCCCAAAC AAGTGTAAG TATTTCCCTA  
TGTGTGTGGA GGGAGGGAGT ATAGGAGGCT GATTTCCTG AGATCCTGTT  
AAATTTGGGT GCCATAGCCA ATCAAAGACG CATCGTTTCC TCTAAGAATT  
CTAAATGGGG CGATTACCAC GGGCCTCGAG GTTCTGGTTT GTATTAGAGG  
AGACACTGTC TTCTTAAGTA AAACATAGAA GGGGAAGTGT CCAGAATTGT  
AAATAAGGCT TCGAGAGAAG CCTTGTCTGG CCACCGGGAT GGAGAAGACC  
TACCTTCGCC TATCCAGGAT CCATCGTCCC TCCCTCTACC CAGATCTGAC  
AGCCCTCCTT GGCTCTTTTG CTGAGGTTTG TTTGAGTTTG TTTTACTCTC  
TGCAAGAGAA GTTCTCTTAA ACATTCTACC CTGTTCAACA GTAAATACAC  
CTCTTAGCTA AGAGGCCACA CACCCAGGGG GAACACCGAT AAAAAGAACA

Figure 1 continued



AGCCAGAACC TTCAGAACGC TGTCGATAGG TACACCAAGC AGCCTTCATA  
CGGAGTTTTC ATTCGTGAGG AGCTGAATAT ACAACAAAGC TAAATGTGAG  
CAGACCAGGC ATGCCTCTGC TAAATGAGGA TGCCACACC AAACATGCCC  
AAGATCTTCA AGTATAATTT TATTATATAG ATTCGCTATG TGTGACATG  
TTTTTATAGT GAACCTGGAT TTTACAAACC CTCTGGTTT GCCACCTGCT  
TCTGGCACCA TACTTGAGGC TTAGGCACGT GATAAAGGAG CATGCCTGTT  
TCCCCCTTA TTTTTTTTAA AGAAAAGCAC CATGTTACAT CATTAATCAT  
GCATATCAGT GTAGTTTAGA TCCGATGTAG AGACAATAAT CTTATCTCTT  
TGTCTGGCTG AAAGACTGTC CTTTAAACTA TCATTCTAAA TGCATTGGT  
TTTTGCCAGG AGTAAACAT GTCACAAGAT ATTTGTTGTC ATTTCCAGG  
-1010 CGTGAAGGA AAGGAATGGA AAGAAAACCA GGGGTGAAG CTGCTGTTCC  
TCTCTAGTCG CTA CTGTAAG TCTACATAGC TGGGGGGGGG GGGGGGACTG  
TTCACATGGG ACCGGTTTCC TCTTGTGTTCC TACACTGGCG CCTCTGGCAA  
AAAACCTCCTC CTTCTCTTCC CCCCAAGCAT ATCTTGGCTG AAAGGTGAGC  
TCTGAAAAGG GGCCTGGCCA AAGTTACTGT AGGGGACCGT GGTGATGGAA  
CTGGGTAAAC AAAAGCACTC TAGCAGCCAC TGGAAAAGGA CCGGGGGCTC  
TTCTCTGTGC ATTTGCCCTG GAACCCCTGAC CACCGCCAGC TCCCTGCATC  
TCCTTGCTAT GGGTTTCTG GACCGACCCA GCCAGGAAGT TCACAACCGA  
AATGTCTTCT AGGGCTAATC AGGTAACCTC GGACGATTTA AAGTTGCCAG  
ATGGACGAGA AAACAGTAGA GGCCTTGGA ACCTGGATAA GCGCCTATCT  
-510 TCTAATTAAA ACATTCAGAC GGGGCGGGG ATGCGGTGGC CAAAGCACCA  
TAAAACAAAA CTTCCAAGTA CTGACCAACT CACTGCAAGT TTGTGCCCCG  
AGTACATCTA GGTTCAGGGG TTCTTGTCTT CATGCTCCCA ACTGCGGGCG  
GATTTTGGT CCCTTGGGAC TTTCAGTGCA GCGGCGAAGA GAGTTCTGCA  
CTTGACGGCT CCTAATGAGG GCGCAGTGGG CCTCGTGTTC CTGGTGATGC  
TTCCAGGTT GCTGGGGGCA GCAAGTGCT CAGAGCCCAT TACTGGCTAC  
ATTTTACTTC CACCAGAAAC CGAGCTGCGT CCAGATTGTC TCTCAGATGC

Figure 1 continued

6/21

GACTTGCCGC CCGGCACAGT TCCGGGGTAG TGGGGGAGTG GCGGTGGGAA  
 ACCGGGAAAC CCAAACCTGG TATCCAGTGG GGGCGTGGC CGGACGCAGG  
 GAGTCCCCAC CCCTCCCGGT AATGACCCCG CCCCCATTCG CTAGTGTGTA  
 +1 (transcription start)  
 -10 GCCGGCGCTC TCTTCTGCC CTGAGTCCTC AGGACCCCAA GAGAGTAAGC  
 TGTGTTTCCT TAGATCGCGC GGACCGCTAC CCGGCAGGAC TGAAAGCCCA  
 GACTGTGTCC CGCAGCCGGG ATAACCTGGC TGACCCGATT CCGCGGACAC  
 CGCTGCAGCC GCGGCTGGAG CCAGGGCGCC GGTGCCCCGC GCTCTCCCCG  
 GTCTTGCGCT GCGGGGGCGC ATACCGCCTC TGTGACTTCT TTGCGGGCCA  
 VRE  
 GGGACGGAGA AGGAGTCTGT GCCTGAGAAC TGGGCTCTGT GCCACGCGG  
 AGGTGCAGGA TGGAGAGCAA GGCCTGCTA GCTGTCGCTC TGTGGTTCTG  
 CGTGGAGACC CGAGCCGCCT CTGTGGGTAA GAAGCCCACT CTTTAGTAGT  
 AAGGCGGAGA AGTAGGGTGC GGGCGGAGAG TGGGAATAGA AGAGGACCTA  
 ACTCGTAGAG CTCTAGAGAC CCTCTCCCT TGGGTGTTCT TCACTTACC  
 +490 AATGGGGA AATGAGTTCA AAGACTCTC CGAATGACT CAGCCAGGAT  
 TCTACTCTCC CCCGGGCATC GGTGGAGCG TGTCTGCGG AGCCGTGACA  
 GCCCTGGCG CTAGGTAGGC AGGAGTGAA AGGCGGCTG AGCCGGGGCA  
 GGAGATGCTC CACTGGCAG GAACAGGCGG TCAACGCTG GGAAGCCAGC  
 TCAAGCCAAG CGGCCCGGCT GGCATCAATC ACTCGGTGCT GTTGCCCAACC  
 GCCCTAGTGG GGGGAGGGA ATCCGCCTCT GGCTCCGCTC CCTTTAGCT  
 CCAGCGTGTA AGCGCACGGA CTATGTGAGG GTAGGTCTCT TCATAGAGCA  
 ACACTTTCCT CCCTCAACTT TCTTTGATGC AGAATGCTAT TTTTGCTGGT  
 AGGAGGAAGA CGCGCTTTC TCTTCTGTGA CAGCTTCTCC AGGTGTATTA  
 AACTAAATAA CTCTCCACTT ACCGACTCCA AAGCGCTGGT CCTGGGGTAA  
 +990 ACTCTGAAAG TCTCAGAAAC TCTTGAGCTT GGCACCTAGT TATAGGTCAC  
 TTTTCTTGTT TTAATAATGCC CTCTGCTTCA AGGTAGGCC CACTCGCT

Figure 1 continued

7/21

CTTGCGCTTT TGTGCAATAA TTTCCCTTCC CTTCCTTCC CTTCCTTCC  
CTTCCTTCC CTTCCTTCC CTTCCTTCC CTTCCTTCC CTTCCTTCC  
TTCTCTCTCC TCTCTCTCT CTATTCTCT GTCATTTCTT TTTTGAAGCC  
ACAGTTTGCA GATTTCAT CTCCACCCAT TGGAGAATGG AGAATCAGGA  
AAAAAGAAGT CAATTCTGCA GAAACATTCC TTGCGCCCTA AGAGAATCGC  
ATGGCTTAAA AGCATTGGCA CTGACATACG GCGCCAAGAT CGCCTGTCTA  
GAGCTATTGA GTTTCTCTCA TAATGACTTG GTTCATCAGG CTAGCTCCAC  
CACGAGTGCC CTCTGTCTCC TGAGAAGGCC GCACTCTCCC CCTTTCTGGG  
AAGAGAAAAG CAGCCTGGAA CATGTGCTTG CCCTGGGTTC CATAGAGAAG  
CAAGTTGCTT TAAAGCCCAG AGAATTCCTA GTGTAGCAGC TTAACAGCGT  
CCCCTCTCT CTGAATAGAT GAGGTTGCC TTTTGGAGTG TGTGACTTGC

XhoI (+1677)

TTAATTGGAT TGGGCTATAA TTGGTGCCAT CCAAGTCTCG AGACAGAGCC  
GCTGTGTGTT TTCCTTCTGG TCTTTGAGCG GGAAGGATAA CAGTGCACAA  
ATTAATTAAT GTTGGTTATC GGATTTGAAC ATAAAAGGCC TTTTATTGTA  
TAGTAGCATA TGTACCTCTT GCAGTCAGAA TGAGCTGTCT AAAGAACAGA  
ACCCAAACTT GCCGATGAAA ATGAATGAGG TTTAATAAAG GCGATGGATG  
AGCATTAGTC ACTGATGTAA ATCTCCAGTT ATTGATAACC TCATTGACTG  
GATTTGATTG CAGACATGTA TTGGTATGGG GCATCCTTTA AAGATGAGCA  
+1990 TAGCCAACGT GCCTGCACTC TAAGAGAATC TATGGCTGTA TGTATTACA  
GAGACAGTTG AGAAGCTCTT AGTGGCTCTG GCGTGTAGAT CAGCGGTAGA  
GCGCTGAGGC TCTGCGCTCG CTTCCTGGCA CTGAAGAATA AAGGCCATTT  
ACTGTGGTGG TGCAGTGGGC GCAGTTTGTG ACGAGTTACT ACTACATTTT  
CCTCACACAT CTGCTGACT AATGAGTTCA TCAGATGAGC GTATCCAGTG  
ATTGTTTGCA GGTTAATGGT TCTCAGTCAT GTTTAGAATC TACTTATCAA  
ACAAATGTTT TTCTCATTTT CTGCTTCTTC TCAAACAAAG TAAGATTCCA  
TTATTGAAAG GCTTGTTTAA GAGCATTTTA ACTGCTTGCC TATGTTAGGG

Figure 1 continued

ACAGTGACTT ATTTTCATATT GACAAATATT ATGCCGATTA ATTGAATATG  
ACTACCCAGT TCTATAGCTG TCTCAGGGCA GACCAAGAGC ATCTGTGATC  
CAGTCACTTT AAATGCCATT TAAATGTCAT AATTGTGTGG TCTAGGAATA  
AACACACTGT AAAGTTTGA ATCACGGCCC AACACAAGT CTTTAACAAT  
GCCAACTAGC TTCTGAGATT CATTAATGTC ATTTAATTAC CAATGTTTTA  
AAAATATGTC ATTAATTACT AAATCTATAG TTGTAACAGC AACACATGTA  
CATCTTATTA AGTTGGGTAT ATTCAGGGTG GCATAGCTGT AGACTATTGC  
ACATCTGTGT TGGTGAGCCA GTGGAGAACT GCCTCCTGGC TGTCTCAGA  
AGGCCACAGT GTCACGGCAT TGGCTATTGT CCTTGGCTCT TTGCTAATAC  
TTTATTGACA TGGCCTCATC TTCGTTACG TTCACTTATT TGCCCAACAA  
CGTCAATGCC AGCTGAGGCC TTAGGAGTCA TCTGTTCTTA GTCAGTGCGA  
ATTAGAAAGC CTGGATGCCT GCCTGCTATT AATTAGTTAT TCTTCTCTTC  
+2990 TGAGACAGAG TCTCACTGTG TGGCCCAGGC TAGTCTCAA CTTCGGGTCC  
ATTTGTCTCA CTCATCAGAA TGCTGGGCTT CCAGGTGTGT GCACCACACT  
AGGTAGCTCG CGTTTAAAGC TAAGAGCTGG AAGATCCTGA TGTCCCTTAC  
CATGGTGGGC ATGTTACAGG TTAGTTGACT GAAACTAGT TATCTCGCTG  
TGTAATGACC TGCAGTGGTA TGTATCTCTC AAGATGCTTT TTTGCATTTC  
AATCAGTTAG GTAACAAGTT CTTAAGTCTC CAGCTTGGA TTGGCATGAG  
CTCAGAGCTT TGATTAATGA GTTGGGACCC CCTAGCTATT GCTCATTAGA  
CTTACACTAT TTTTAGTTTT GCTCTGAGTT TATGAATATG CATGTATGCA  
TGAACITGGG AGATATTTTT CTTCCTCCAA TCCTTTTCTT CCATTAAAT  
GTGCTGTCTT TAGAAGCCAC TGCCTCAGCT TCTGCAGCTC AGATACCAAA  
GGAAGTCTGG TACACAGCAT GATAAAAGAC AATGGGACGG GGTACAGTG  
GCTCCCGTCC CTTTCAGGGG TATGGAGACG AGCTGTAGAG AGATGTCTCC  
AGGGAGTTTT CATTAATCAG CAATTTAGTC AGATCTGTGC ATCCTATGCT  
TTACAAGAAA TGTCAGTGGG CCTGAGATCA TCAGATGGAG GTTCATCGGG  
TTTCAATGTC CCGTATCCTT TTGTAAGACC TTGAAGTTGG CAACGCAGGA

Figure 1 continued

AAACAGGAAC TCCACCCTGG TGCCGTGAAT TGCAGAGCTG TTGTGTTGGT  
 TTGTGACCAT CTGCCCATTC TTCCTGTTAT GACAGAGCTT GTGAACTTTA  
 ACTGGGACTG GGGCAAAGTC AATCCACCT TTATACAATG AATTGCTGAA  
 GAGGCCTTTT AAAACTTGGA GTGTGCATTG TTTATGGAAG GGCTTTCTTA

BamHI (+3947)

+3990 TTGGATCCAA CTCTTTTCTA ATTTGTTTCT AGGTTTGCCT **GGCGATTTC**  
**TCCATCCCC CAAGCTCAGC ACACAGAAAG ACATACTGAC AATTTTGGCA**  
**AATACAACCC TTCAGATTAC TTGCAGGTAA** GGATTCCTTT TTGAGCCAGC  
 TTTCTATGT GAAAGGACTC ATTGTTTACT GAGGTCACAA CAATTTCCAC  
 TATTGCAGAA GTATAATAGT ATTGTTACAA TTGTTTATAA ATCATGAGAC  
 TTCTAAGAAC CTATTTAATA ATGAAACAAT GGAAAAAGTC TTTTCAAACC  
 TTTGTACTCT TTTGCTGAGC CGTTTTCAAC ATGCACAAAC ATATTACACA  
 AATATAACAT ACACAGGAAC ACACATGAAT GCATGGGATG ATGTGCCATAA  
 AACTAGCATG TAATTGATAT TCACAATTAT TGATAAATTA GTAAAGCAAA  
 GGAATTCCTT ATGAATAGAG CTAAAATTCT ATCCATGTTC AAGTCACCCA  
 GAATGGCTTC TGGACATTTT TTTTTTTAGC TGTTTTCTAC AAGTGAAATT  
 CTGCCTGTAT TAGCAATTTA ATATCTAGCC AATAATATTC CTGACCATAT  
 GTCCTGTTC AACCATGACC TTCATAATCT GGCTTGATGT TCTGGGCTTC  
 TTTCCCTCTT GCCAGCAAGA TGTACCGGTG TTGATGCTGG ATAAACTGAG  
 AAACAGAAGT TTTTCGCAAG AAGAGGACCT TGAATTTTGC TTTTCCCCTG  
 AGAGACAAGA AAGGAACTT AGAGGAGGTG TAGCTGGGAG TGTGGTCATT  
 CATGAAAGAC CTGTTTGCAG GGCAGTGTGT TTTGCTGGGG ACAGTAATGA  
 GCCTAGATCG TAGTGCCATC CCAAGAGAGT GCTTGGTGGC AAAAAAGGCC  
 CTAGCAGCTT GTGGCAGTTG CCTCATATTT GAAGAATACT AAGAGGTCCC  
 CCGAATAACT CAGGGCTAGT GTTGATCATT GCATGTGGAG AGAATCCAAG  
 CCTCCTATCT AGGGTCTACA AAAGTAACCA ATGCCAGTC TTTGGGGGAA

Figure 1 continued

10/21

+4990 AGCAAAACCA GAAAGCGATG ATAGCAGGAC CTGTTTATTT TCATTAAGTC  
 ATGGCATTTC CAGAGACTTT GCTCCCCCTA TTCTCAGACA CAAAGCCCAC  
 TTAAGATCTC CCTCTGGAGA CTGCTGGGAA CATTTCTTAA GTTCTGAAAA  
 AACCCTGGAG TGATTGGGCA CAGACGATCC TGCACTTCA TGTGAGTGTCT  
 AAGCTCTTTG GGTGATGACT CAGTGGGTCA CATTTGTTTA TTCATATTGA  
 CTACCTTCCG TTTGCTTTGC GGAGAATGGA AGCTATAGAA GTCTGTTTGG  
 TGTGGCCCTC ACAAGGCACT GTGAGCTTCT TCTCTCTGTG TGCTAACTTC  
 TTACTCTCCC TTGCTTATAC CCACATAGGG ACTCTGGCTT TGTGTCTGTT  
 CTTCAATGCT TCAGATGTGC CCTGGGTCTT GTCTGTCCTT CACACTTACT  
 GATGCTGCCT GGAATGCTAT TCCTCCCAAT GTGCATAGGG CCAGCTCGGT  
 CCAAACTCTC TCTTTTCTTT GCCTCTTTTA TATTTTCTT CACAGTATCA  
 AATCACCACA GTTTATGCAA CAAACTGAAA CTTTAAATTT GTCTGTCTCC  
 TTATATTAGT GATAGGTTCC AGAAAGGCAC TGATTTTTTT TCTTCCTTGG  
 TGTACACTGG GCAACTACTC TACCACTGAG CGTGATATCC TTGGTCCCTT  
 AAAAGTTATC CTCTGTCCTT AATAATGCTT AGCAATCATA TTTGCTTAAA  
 ATATTATTATG AATGACTGCA GGAATGAATG AATGAATGAG CTAACAGAAA  
 ACTCATGACC ATGTGGGTGA TTTCCGAAAC AGAGTGTGAG ATCTTTGGTG  
 GCATGTCCTT GTAGACTGTC TGCCACCAGT ATCTATCATC TTGAAGGTGA  
 CTATTGAGTA GTTTATATGC ATGTGAAAAA CCAAACTTTC TATTCTCTTA  
 CTCATAGCCT CTCTTAATCA TAGCCCTGTG GCATGGAGTG TACCATTGAT  
 +5990 ATCTTCCTGG AATACTTTTT CAGGGGACAG CGGGACCTGG ACTGGCTTTG  
 GCCCAATGCT CAGCGTGATT CTGAGGAAAG GGTATTGGTG ACTGAATGCG  
 GCGGTGGTGA CAGTATCTTC TGCAAAACAC TCACCATTCC CAGGGTGGTT  
 GGAAATGATA CTGGAGCCTA CAAGTGCTCG TACCGGGACG TCGAC  
 (SEQ ID NO:1)

Figure 1 continued

11/21

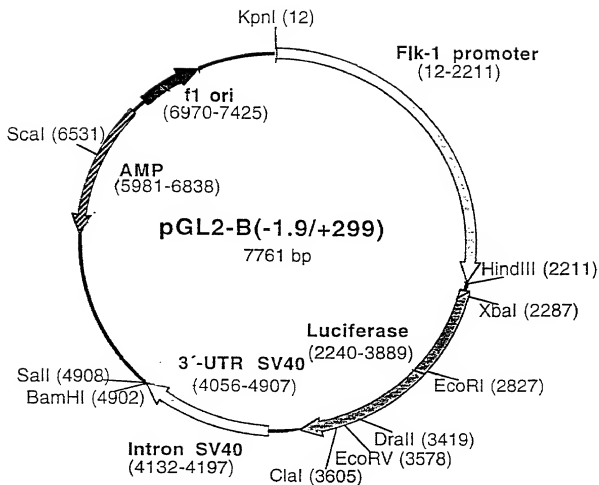


Figure 2

12/21

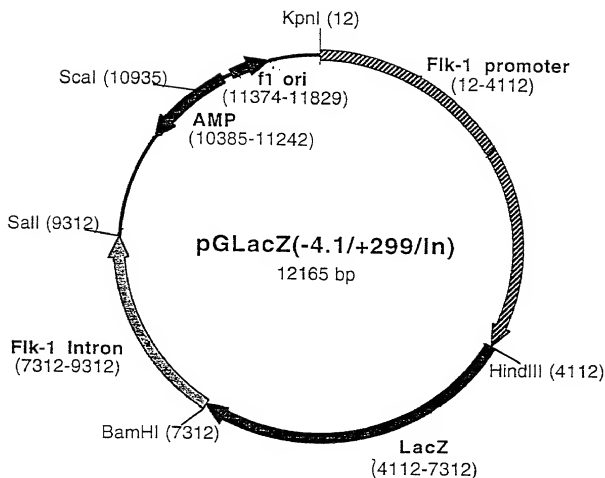
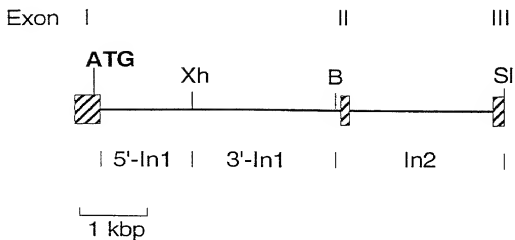


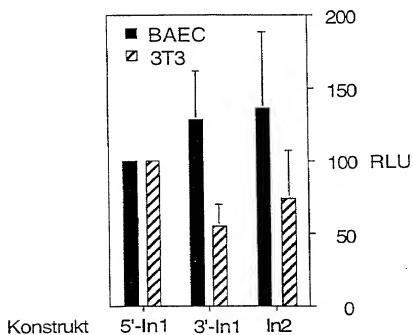
Figure 3



13/21



A



B

Figure 4

14/21

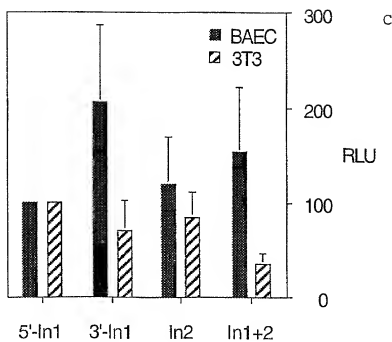


Figure 4 continued

A



B



Figure 5

15/21

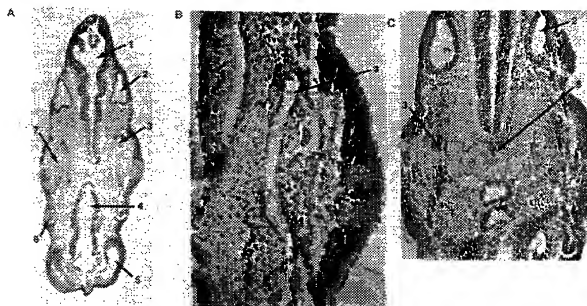


Figure 6

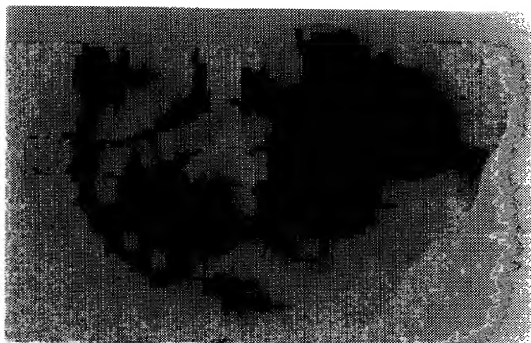


Figure 7

16/21

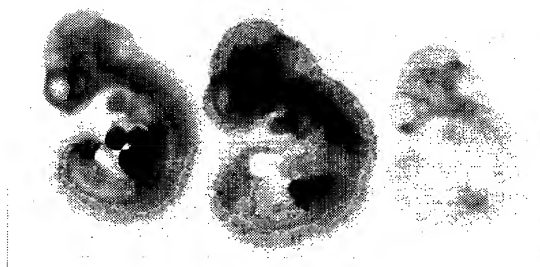


Figure 8

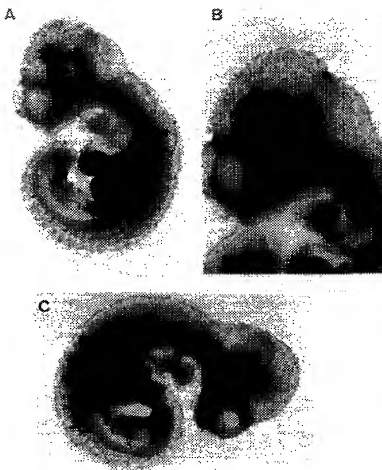


Figure 9

17/21

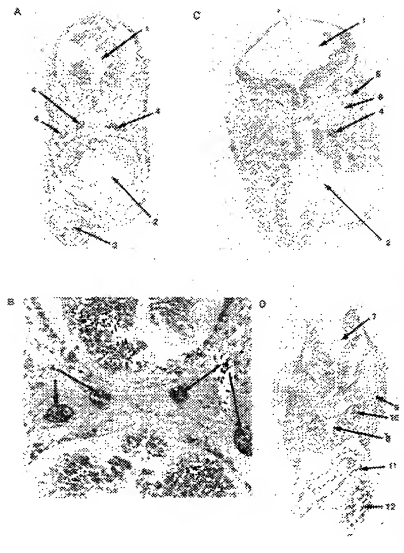


Figure 10

18/21

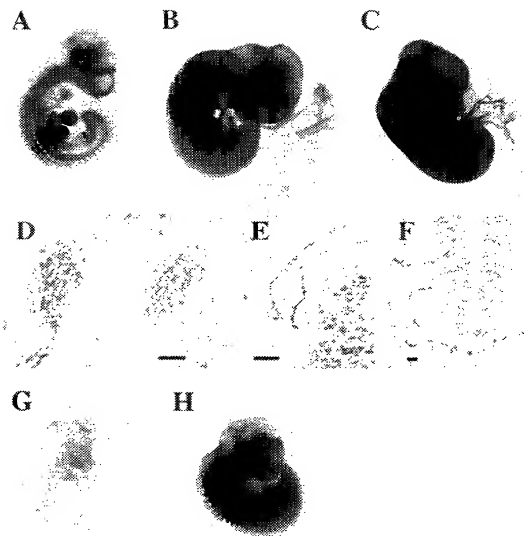


Figure 11

19/21

GATA PEA3

AAATGTGCTGTCTTTAGAAAGCCACTGCCCTCAGCTTCTGCAGCTCAGATACCAAAGGAAGTCTGGT

GATA AP1

ACACAGCATGATAAAAGACAATGGGAACGGGGTCACAGTGGGTCCCGTCCCTTTTCAGGGGTATGGA

NFkB AP1

GACGAGCTGTAGAGAGATGTCTCCAGGAGTTTTCATTAATCAGCAATTTAGTCAGATCTGTGCA

STAT SCL/TAL-1

TCCTATGCTTTACAGAAATGTCACTGGGCCTGAGATCATCAGATGGAGGTTCATCGGGTTTCAA

Ets-1 GATA Ets-1

TGTCCTGATCCCTTTTGTAAGACCTTGAAGTTGGCAACGCAGGAAAAACAGGAACCTCCACCTGCT

SCL/TAL-1 Ets-1

GCCGTGAATTGCAGAGCTGTGTGTGGTTTGTGACCATCTGCCCATCTCTCTGTATATGACAGA

GCTTGTGAACTTTAACTGGGACTGGGGCAAAGTCAATCCACCTTTATACAATGAATTGCTGAAG

AGGCCTTTTAAACCTTGGAGTGTGCATGTGTTATGGAAGGCTTTCCTATTGGATC

Figure 12

20/21

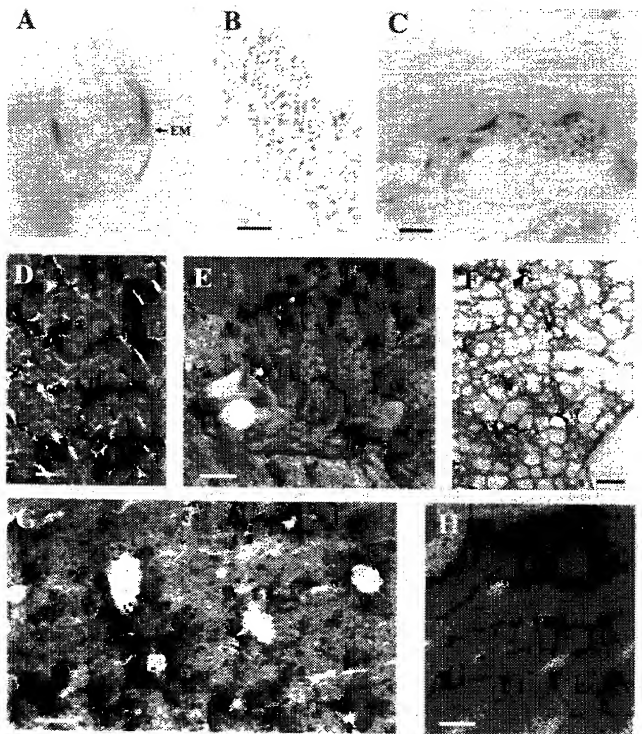


Figure 13



21/21

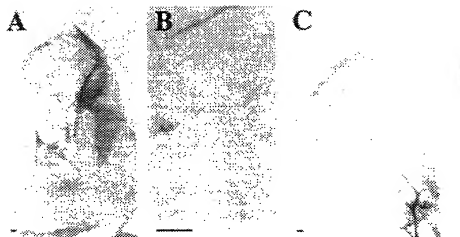


Figure 14

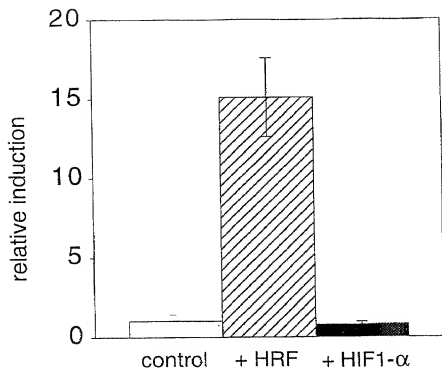


Figure 15

PATENT  
Attorney Docket No.: VOSS1110

In re Application of: )  
Breier et al. )  
Application No.: 09/445,201 )  
Filed: December 2, 1999 )  
For: REGULATORY SEQUENCES )  
CAPABLE OF CONFERRING )  
EXPRESSION OF A )  
HETEROLOGOUS DNA SEQUENCE )  
IN ENDOTHELIAL CELLS IN VIVO )  
AND USES THEREOF )

POWER OF ATTORNEY BY ASSIGNEE

As a below-named assignee of the above-identified application ("Application"):

I hereby appoint the following attorneys of the assignee to prosecute the Application and to transact all business in the United States Patent and Trademark Office connected therewith:

TIM ELLIS  
LISA A. HAILE  
WILLIAM N. HULSEY, III  
RICHARD J. IMBRA  
SHEILA R. KIRSCHENBAUM  
JUNE M. LEARN  
TIMOTHY W. LOHSE  
TERRANCE A. MEADOR  
JOHN OSKOREP  
STEPHEN E. REITER  
STEVEN R. SPRINKLE  
RAMSEY R. STEWART  
DAVID R. STEVENS  
BARRY N. YOUNG

Registration No. 41,734  
Registration No. 38,347  
Registration No. 33,402  
Registration No. 37,643  
Registration No. 44,835  
Registration No. 31,238  
Registration No. 35,255  
Registration No. 30,298  
Registration No. 41,234  
Registration No. 31,192  
Registration No. 40,825  
Registration No. 38,322  
Registration No. 38,626  
Registration No. 27,774

In re Application of:  
Breier et al.  
Application No.: 09/445,201  
Filed: December 2, 1999  
Page 2

PATENT  
Attorney Docket No.: VOSS1110

I hereby authorize and request insertion of the application number of the  
Application when officially known.

Direct all telephone calls to:

LISA A. HAILE, PH.D.  
Telephone: (858) 677-1456

Address all correspondence to:

LISA A. HAILE, PH.D.  
GRAY CARY WARE & FREIDENRICH LLP  
4365 Executive Drive, Suite 1600  
San Diego, CA 92121

**Max-Planck-Gesellschaft zur Forderung der Wissenschaften**

By: Christa Herzog  
Name: Christa Herzog  
Title: Head of patent department  
Date: 15.3.2000

other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

COUNTRY	APPLICATION NO.	FILING DATE	PRIORITY CLAIMED
<u>Germany</u>	<u>EP 97 10 8959.4</u>	<u>6/3/97</u>	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
_____	_____	_____	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of first inventor: Georg Breier

Inventor's signature: *Georg Breier*

Date: March 17, 2000

Residence: Eibenweg 1, 61231 Bad Nauheim, Germany *DEX*

Citizenship: Germany

Post Office Address: Eibenweg, 1, 61231 Bad Nauheim, Germany

Full name of second inventor: Werner Risau (deceased)

signed by Barbara Risau as executor of his estate

*2-01*  
Inventor's signature: *Barbara Risau*

Date: *18. 03. 2000*

Residence: Dresdner Str. 2, 35510 Butzbach, Germany *DEX*

Citizenship: Germany

Post Office Address: Dresdner Str. 2, 35510 Butzbach, Germany

Full name of third inventor: Volker Rönicke

Inventor's signature: [Signature]

Date: 13 III 00

Residence: Danzigerstr. 18, 67240 Bobenheim-Roxheim  
Cottbuser Str. 6, 35415 Pohlheim, Germany Danzigerstr. 18, 67240 Bobenheim-Roxheim, Germany

Citizenship: Germany

Post Office Address: Danzigerstr. 18, 67240 Bobenheim-Roxheim  
Cottbuser Str. 6, 35415 Pohlheim, Germany

as of May 2000

(effective as of May 2000:

Neufriedenheimerstr. 58a

Neufriedenheimerstr. 58a  
81375 München, Germany)

München / Neukirchen

01212

DECLARATION FOR PATENT APPLICATION

As a below-named inventor, I hereby declare that:

My residence, post office address and citizenship is as stated below next to my name.

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled **REGULATORY SEQUENCES CAPABLE OF CONFERRING EXPRESSION OF A HETEROLOGOUS DNA SEQUENCE IN ENDOTHELIAL CELLS IN VIVO AND USES THEREOF**, the specification of which

\_\_\_\_\_ is attached hereto.

  X   was filed on December 2, 1999 (Attorney Docket No. VOSS1110)  
as U.S. Application Serial No. 09/445,201  
and was amended on \_\_\_\_\_  
if applicable (the "Application").

**I hereby authorize and request insertion of the application serial number of the Application when officially known.**

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability of the subject matter of the Application as defined in Title 37, Code of Federal Regulations ("C.F.R."), § 1.56.

With respect to the Application, I hereby claim the benefit under 35 U.S.C. Section 119(e) of any United States provisional application(s) listed below:

<u>                    </u> (Application Serial No.)	<u>                    </u> (Filing Date)
<u>                    </u> (Application Serial No.)	<u>                    </u> (Filing Date)
<u>                    </u> (Application Serial No.)	<u>                    </u> (Filing Date)

With respect to the Application, I hereby claim the benefit under 35 U.S.C. Section 120 of any United States application(s), or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of the application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability of the subject matter of the Application as defined in Title 37, C.F.R., Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of the Application:

<u>PCT/EP98/03318</u> (Application Serial No.)	<u>6/3/98</u> (Filing Date)	<u>Pending</u> (Status) (patented, pending, abandoned)
<u>                    </u> (Application Serial No.)	<u>                    </u> (Filing Date)	<u>                    </u> (Status) (patented, pending, abandoned)
<u>                    </u> (Application Serial No.)	<u>                    </u> (Filing Date)	<u>                    </u> (Status) (patented, pending, abandoned)

I hereby claim foreign priority benefits under Title 35, United States Code, § 119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application for patent or inventor's certificate or any PCT international application(s) designating at least one country